Polymeric neoglycoconjugates: Synthesis, characterization and properties

Ambrosi, Moira

POLYMERIC NEOGLYCOCONJUGATES: SYNTHESIS, CHARACTERISATION AND PROPERTIES

Moira Ambrosi

A thesis submitted to

University of Durham
Department of Chemistry

in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2002

STATEMENT OF COPYRIGHT

DECLARATION

The work reported in this thesis was carried out in the Department of Chemistry at the University of Durham between October 1999 and September 2002. All the work was carried out by the author unless otherwise stated, and has not previously been submitted for a degree at this or any other university.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to the people who have been involved in the completion of this work and without whom I would not have been able to fully achieve the objectives of this study.

First of all, sincere thanks go to my supervisor, Dr. N. R. Cameron, who, with his constant help, encouragement and advice, made these three years an exquisite occasion of growth from a scientific and human point of view. Thanks to him I have had the opportunity to work on a novel and exciting subject to which I hope to be able to continue to give my contribution in the future.

In the same way, I would like to thank Dr. B. G. Davis (University of Oxford), who constantly followed this work with enthusiasm, enriching it with his precious suggestions and hinting on carbohydrate synthesis and biology.

Profound thanks go to Dr. S. J. Cooper, for giving me the opportunity to use her equipments and for sharing with me her knowledge about physical chemistry in general and nucleation/crystallisation studies in particular.

I would like to express my grateful thanks to Dr. S. Stolnik (University of Nottingham), working with whom has been a real pleasure. She gave me the opportunity to spend two months in her lab, using the microcalorimeter. She has been constantly present, following my work daily with profound interest and giving fundamental suggestions for the design of the experiments and for the interpretation of the results obtained.

I would like to sincerely thank Prof. R. W. Richards and Prof. D. Parker, for having kindly placed at my disposal their instrumentations and for the useful discussions.

Particular thanks go to Dr. A. M. Kenwright for his precious hinting for the interpretation of the NMR spectra and to Ms. C. Heffernan and Mr. I. McKeag for performing them.

Thanks to Dr. M. Jones and Ms. L. Turner for the mass spectroscopy service; to Dr. A. Batsanov for performing the X-ray crystallographic measurements; to Ms. J. Dostal for the elemental analyses and to Mr. D. Carswell for performing SEC measurements in THF.

My gratitude also goes to Dr. R. Hunter (Unilever Research, Port Sunlight) for the SEC determinations in aqueous media and to Dr. A. Miller (University of Cambridge) for the ESEM analyses.

Finally, thanks to all those persons who are for me special. Their love and trust has been and is an inextinguishable source of motivation and strength.

Freedom is of whom thinks differently

To my father

ABSTRACT

As part of a currently intense research effort into the development of novel potent macromolecular pharmaceuticals, diagnostic tools and carriers for targeted drug delivery, this project involved the synthesis of glycopolymers. Polymethacrylate derivatives carrying carbohydrate residues were efficiently prepared in a stereo-controlled manner, following two different routes. As revealed by a thorough characterisation of the materials obtained, the synthetic procedure strongly affected the polymers' composition and, in turn, their properties, such as their self-association in aqueous media and the interaction of the compounds bearing D-galactose units with a specific lectin.

Polymers containing peracetylated saccharide residues were shown to be surface-active, forming stable spread Langmuir films. These monolayers acted as effective dl-aspartic acid nucleation promoters, whilst the soluble analogues (for which the sugars are deacetylated) seemed to possess a certain activity as dl-aspartic acid growth inhibitors.

The behaviour of the deprotected polymers in aqueous solution was thoroughly investigated, prior to testing whether the materials carrying galactosyl units on the side chains could be recognised by a galactose-specific protein. For the first time for polymeric glycoconjugates, thermodynamic binding parameters relative to this protein-carbohydrate interaction were directly evaluated using isothermal titration microcalorimetry.

Moira Ambrosi, Durham, September 2002

TABLE OF CONTENTS

1

INTRODUCTION

1.1 LECTINS

Lectins[1] exist in most living organisms, ranging from viruses and bacteria to plants and animals. They were first identified before the turn of last century as plant proteins capable of agglutination of red blood cells[2,3]. Despite this early discovery, it was only in the 1960s that their study started to attract great attention[4,5]. Their involvement in diverse biological processes[6], such as clearance of glycoproteins from the circulatory system, trafficking, adhesion of infectious agents to host cells, recruitment of leukocytes to inflammatory sites, cell interactions in the immune system, in malignancy and metastasis, was then shown. At present they are the focus of intense research. The investigation of their role in cell recognition, as well as their employment as invaluable tools for the study of complex carbohydrates in solution and on the cell surface, is contributing markedly to the advancement in glycobiology.

Like enzymes and antibodies, lectins interact with carbohydrates non-covalently; however, unlike these biomolecules, they do not present any catalytic activity and they are not products of an immune response. The interaction with sugars is usually reversible and highly specific[7,8]. Each lectin contains typically two or more carbohydrate-binding sites; therefore, their interaction with sugars on the surface of erythrocytes results in the cross-linking of several blood cells and their subsequent precipitation. This phenomenon, known as cell agglutination, is a major attribute of lectins' activity and is used routinely for their detection and characterisation. Both the

agglutination and precipitation processes are inhibited by the carbohydrate for which the lectin is specific.

Lectins represent a heterogeneous group of oligomeric proteins, widely varying in size and structure. Nonetheless, many of them can be grouped in families characterised by similar sequences and structural organisation. Sequence similarity with known lectins provides a novel guideline for the identification of new ones.

According to the monosaccharide toward which they exhibit the highest affinity, lectins can be classified into five groups: mannose, galactose/*N*-acetylgalactosamine, *N*-acetylglucosamine, fucose and *N*-acetylneuraminic acid[9]. With the exception of fucose, all the sugars are in the D configuration. Despite the weak intrinsic protein-ligand affinity, with typical association constants in the millimolar range for monosaccharides, lectin selectivity is usually high[10]. Therefore, apart from few exceptions, neither do lectins specific for galactose bind glucose, nor those specific for mannose bind galactose. However, variations at the C-2 position of the pyranose ring may be tolerated, so that most lectins that bind galactose also interact with *N*-acetylgalactosamine. Certain lectins combine preferentially with either the α- or the β-anomer, whereas others lack anomeric specificity. The properties of the aglycon can influence markedly the lectin recognition activity. For instance, aromatic glycosides bind to many lectins much more strongly than aliphatic ones, indicating the presence of a hydrophobic region in the proximity of the carbohydrate-binding site[11].

Besides that based on carbohydrate specificity, other classifications of lectins have been drawn. The determination of the amino acid sequences of several hundreds of lectins, together with the elucidation of about thirty 3D-structures, allowed the replacement of the traditional division according to the lectin origin – *i.e.* plant,

animals and microorganisms – by a classification based on common structural features[9]. Most lectins belong to three classes: (1) simple, (2) mosaic (or multidomain) and (3) macromolecular assemblies. Simple lectins consist of a small number of subunits, not necessarily identical, each of molecular weight usually below 40 kDa. Each monomeric unit contains a carbohydrate-binding site. This class comprises practically all known plant lectins[12] and the galectins (formerly known as S-lectins), a family of galactose specific animal lectins[13]. Mosaic lectins are composite molecules consisting of several kinds of protein domains, only one of which possesses a carbohydrate-binding site. This class includes diverse proteins from different sources: viral hemagglutinins[14] and animal lectins of C-, P- and I-types[15]. Macromolecular assemblies are common in bacteria. They are filamentous organelles consisting of helically arranged subunits (pilins) assembled in a well-defined order[16]. Within each class, proteins can be grouped into families, with similar sequences and structural properties.

In this study we will concentrate our attention on legume lectins, a family of simple plant proteins, including the lectin from *Arachis Hypogaea* (peanut agglutinin, PNA), used here as a galactose-specific protein to test the lectin-ligand affinity of our glycopolymers.

1.1.1 LEGUME LECTINS

Legume lectins represent the largest and most thoroughly studied family of simple lectins. Around 100 members have been characterised, almost all isolated from the seeds of the plants from which they come[12,17]. Concanavalin A (Con A), the lectin from the jack bean, is the prototype member of the family. Firstly isolated in 1919 by Sumner, it was only in 1936 that Sumner and Howell showed its specificity

toward mannose and glucose[18]. The relative abundance of this protein in jack bean, the ease of its preparation and the large number of saccharides with which it can interact, led to numerous studies on Con A, markedly accelerated by the discovery that cells transformed by DNA tumour viruses or carcinogens were agglutinated by the lectin more readily than normal cells. About 85% of the binding sites for Con A, which are in a cryptic form on normal cells, were found to be exposed on the surface membrane of transformed cells. It was therefore hypothesised that the change in structure of the surface membrane, resulting in the exposure of the sites, could produce the change in cellular regulatory mechanisms associated with transformation[19].

1.1.1.1 STRUCTURAL FEATURES

Typically, legume lectins consist of two or four identical or almost identical subunits (protomers) of 25-30 kDa each, in turn commonly constituted by a single polypeptide chain of about 250 amino acids that may present one or two *N*-linked oligosaccharides. Each protomer contains a carbohydrate-combining site, a tightly bound Ca^{2+} and a transition metal ion, usually Mn^{2+}. 20% of the amino acid residues are invariant in all legume lectins and another 20% are similar in all proteins. The conserved amino acids include several of those involved in the interaction with the saccharide and almost all the residues that coordinate the metal ions. The resolution of 3D-structures of about ten legume lectins has shown that each subunit is constituted largely – nearly 60% – of β-strands mutually connected by loops. For all legume lectins known so far, the tertiary structure is made up of two anti-parallel β-sheets, a six-stranded flat "back" and a seven-stranded curved "front", connected by a five-

stranded β-sheet, giving the well known "jellyroll" motif, also referred to as the "lectin fold"[20] (Figure 1a).

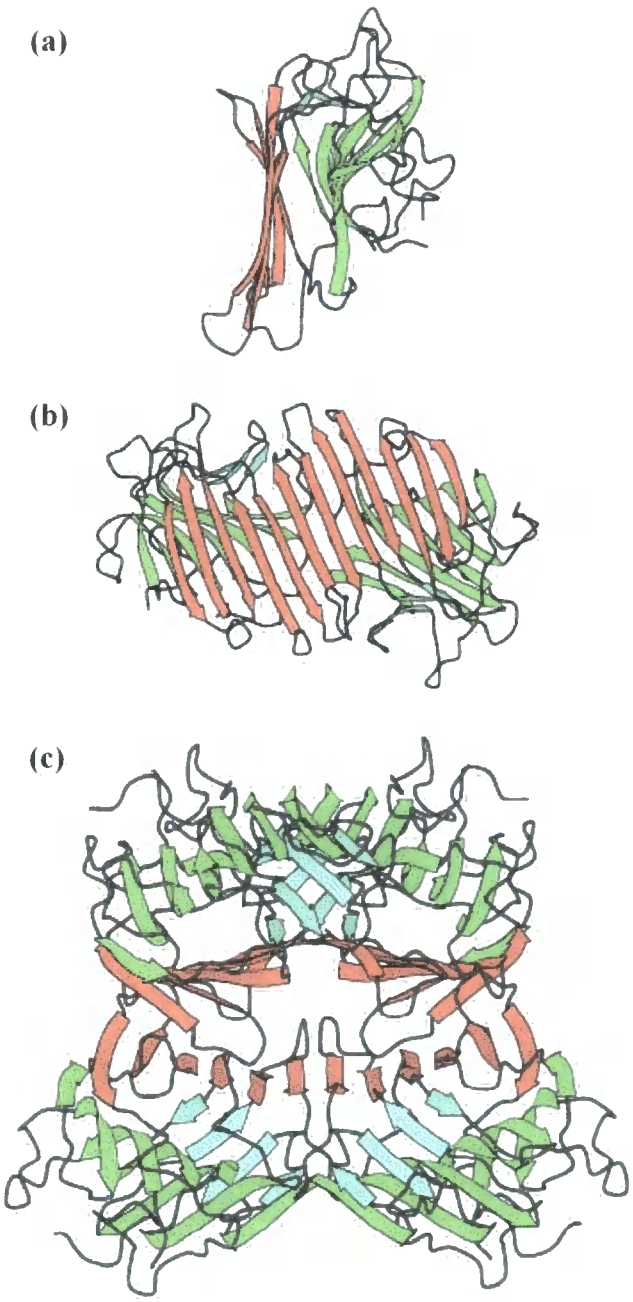


Figure 1 a), Tertiary structure of a legume lectin monomer; b), dimerisation in Con A; c), tetramerisation in Con A.

The subunit structures of different lectins can be nearly superimposed, irrespective of the protein specificity. Despite their similarities at the primary, secondary and tertiary structural monomeric level, legume lectins exhibit considerable variation in their

quaternary structure: small differences in the amino acid sequences at the monomer-monomer interfaces and the presence/absence of glycosylation affect the monomers' association mode. In the case of lectins with "canonical" quaternary structure, such as Con A, dimerisation involves anti-parallel side-by-side alignment of the flat six-stranded β-sheets of the two monomers, resulting in the formation of a continuous 12-stranded sheet that extends across the dimer interface. A considerable portion of the surface area is buried in the process: *ca.* 1000 $Å^2$ per monomer (Figure 1b). Further association of two dimers gives the tetrameric assembly of Con A observed in physiological conditions[21] (Figure 1c).

1.1.1.2 CARBOHYDRATE-BINDING SITE

During the past 15 years there has been significant progress in elucidating the features of lectins involved in carbohydrate binding. Mainly X-ray crystallography of the proteins complexed with their ligands, site-directed mutagenesis experiments and molecular modelling have allowed the identification of the chemical groups belonging to both interacting species involved in the binding and of the types of bond formed. Studies of lectin-oligosaccharide complexes are especially interesting, providing the basis for the understanding of the proteins' interaction with natural ligands. Generally, lectins show exquisite specificity for di-, tri- and tetrasaccharides, with association constants significantly higher compared to the corresponding monosaccharides.

The carbohydrate-binding site is represented by a shallow depression on the surface of the protein, located at the top right of each protomer. In all cases the combining site appears to be preformed[22], since few conformational changes occur upon binding. The Ca^{2+} and Mn^{2+} are situated around 4 Å apart and in close proximity to the sugar-combining pocket. Although not always directly involved in the

carbohydrate binding, the cations help the positioning of the amino acid residues interacting with the glycoside.

Lectins bind carbohydrates through a network of hydrogen bonds and hydrophobic interactions. A schematic representation of the T-antigenic disaccharide, Galβ(1→3)GalNAc (Tant), in the binding site of PNA is reported in Figure 2.

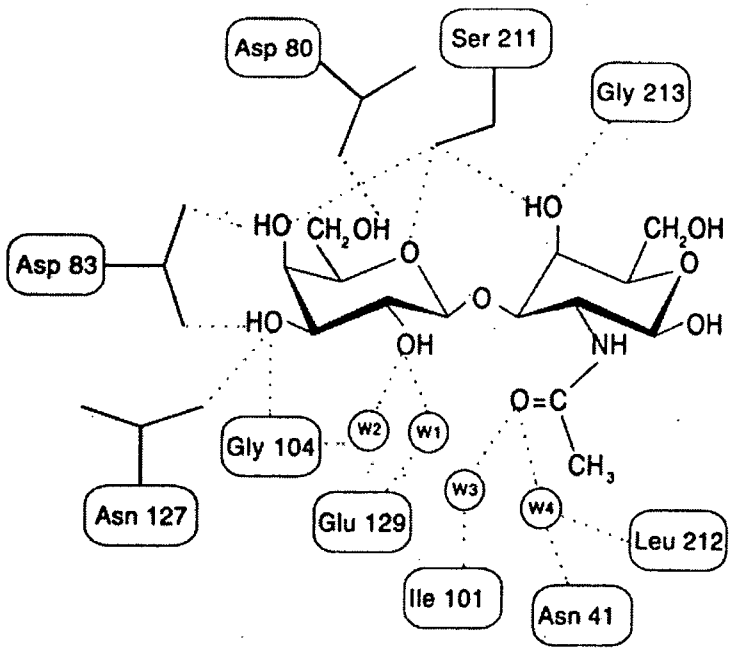


Figure 2 Schematic representation of protein-carbohydrate interactions in the PNA-Tant complex.

Van der Waals forces, although rather weak (usually a fraction of 4.2 kJ mol^{-1} for each pair of atoms), are frequently numerous, contributing significantly to the overall binding[9]. The steric disposition of hydroxyl groups creates hydrophobic patches[23] on the sugar surface that can interact with hydrophobic regions of the protein. A common type of interaction is the stacking of the sugar on the side chain of aromatic amino acids such as phenylalanine, tyrosine or tryptophan[24]. Contacts between the ligand and the protein are often mediated by water molecules. Water acts as a molecular "mortar"[25]; its small size and its ability to behave as both hydrogen donor and acceptor make it ideal to exploit this function. Tightly bound water

molecules can be considered as structural, *i.e.* an extension of the protein surface. Thus, water plays a significant role in carbohydrate recognition, imparting in some cases exquisite specificity*.

In all legume lectins, irrespective of their specificity, four invariant amino acid residues participate in the ligand binding: an aspartic acid, an asparagine, a glycine (conserved in all the lectins of the family apart from Con A) and an aromatic amino acid[26] or leucine[27]. Replacement of the aspartic acid or asparagine by site-directed mutagenesis results in several cases in the loss of the lectin sugar-binding ability. The aspartic acid and asparagine residues also participate in coordinating Ca^{2+} (Figure 3).

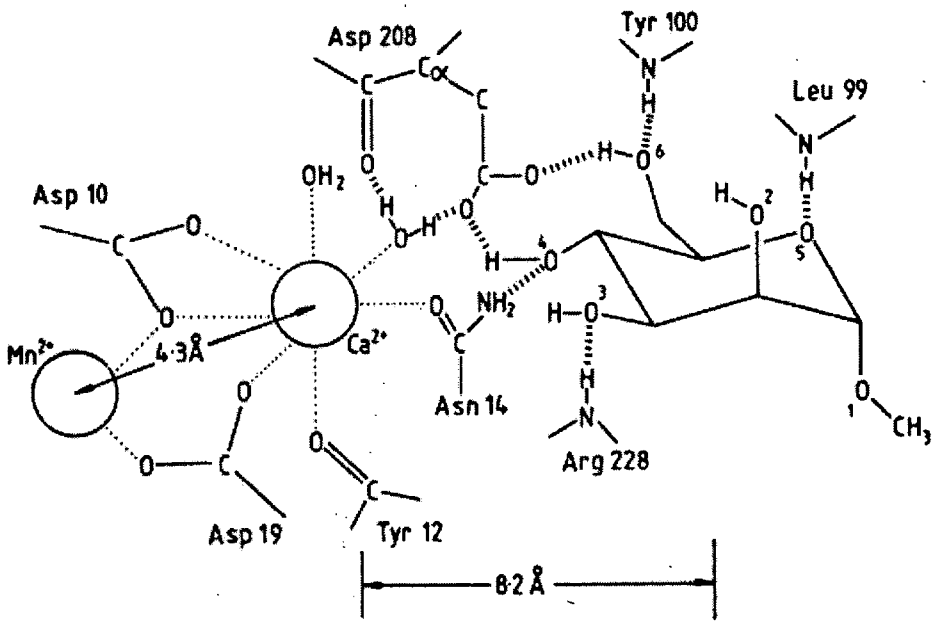


Figure 3 Methyl-α-mannoside in the combining site of concanavalin A.

A rare *cis*-peptide bond between the critical asparagine and the preceding amino acid, usually alanine, confer to the asparagine residue the proper orientation.

* As will be discussed in Chapter 4, the high specificity of peanut agglutinin for Galβ(1→3)GalNAc appears to be generated by the formation of two water bridges (the involved water molecules are labelled as W3 and W4 in Figure 2) absent in the binding of other ligands.

Despite the conservation of key amino acids involved in the binding of the carbohydrate, different legume lectins can show different specificity. For instance, while Con A binds mannose and glucose, PNA, ECorL and SBA bind galactose. As reported in Figures 2 and 3 and better visualised in Figure 4, the aspartic acid and the asparagine residues form hydrogen bonds with different hydroxyl groups of mannose and galactose, due to the different orientation of the monosaccharides in the binding sites of the respective specific lectins[28].

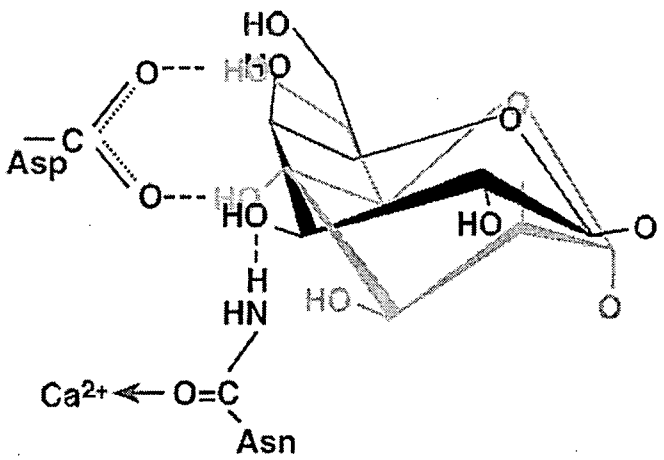


Figure 4 Hydrogen bonds between the side chains of conserved amino acids that bind mannose (lighter lines) to concanavalin A or *Lathyrus ochrus* lectin I and galactose (darker lines) to ECorL, or PNA, or GSIV.

Therefore, while the constellation of highly conserved amino acids provides the framework required for binding, specificity apparently arises from the variability of amino acids residues in other regions of the combining pocket. The amino acids that form the binding sites of legume lectins lie in four loops of the protomer polypeptide chain, designated A, B, C and D[29]. The invariant aspartic acid and glycine belong to A and B, respectively, whereas the asparagine and the hydrophobic residue are in loop C. Additional interactions are provided by amino acids in loop D, indicating the correlation between this loop and the lectin specificity. Thus, for instance, the size of loop D is identical in all mannose-specific lectins[30].

1.1.1.3 FUNCTIONS

Although known longer than lectins of different origin, the physiological role of legume lectins is still not well understood. Many theories have been proposed in the course of the years, among which only two have received serious attention.

According to the first one, they are involved in the establishment of symbioses between nitrogen-fixing bacteria and plants[6,31]. Legumes are able to associate specifically and form symbioses with soil bacteria of the rhizobia family, a phenomenon that makes them independent from soil nitrogen supplies. The idea that lectins are responsible for this association was advanced over 20 years ago. Nonetheless, the mechanism of the process is still obscure. The finding that a legume binds to a specific rhizobial species and not to bacteria that are symbionts for other legumes suggests that the interaction occurs between lectins in the roots of the plant and carbohydrates on the bacteria surface. Despite the fact that molecular genetics experiments support this hypothesis, several inconsistencies can be pointed out. First of all, for most host-symbiont systems there is no proof of the presence of lectins and of the respective ligands on the two interacting species. Secondly, the correlation between the sugar specificity of legume lectins and their ability to recognise bacteria appears not to be particularly strong. Furthermore, mutants of soybean lacking the lectin are still nodulated by the rhizobial symbiont[9].

The second theory is more general, embracing all plant lectins. In this assumption, lectins are considered as defence agents against predators[32,33]. An essential feature of any active defence agent is the ability to specifically recognise the pathogen. Early investigators noted the similarities of lectins to antibodies and hypothesised that lectins might function as plant antibodies[2,3]. The toxicity of several plant lectins for animals and their inhibition of fungi growth are well known

phenomena. Nevertheless, unlike antibodies, lectins are not produced in response to the exposure to an antigen. Moreover, information about the toxic effect is scarce and mainly the result of feeding experiments using *Phaseolus vulgaris* (PHA) and of accidental poisoning of humans. Feeding experiments carried out with some animals showed their preference for lectin-free plants and their reluctance in consuming food containing PHA.

1.2 CARBOHYDRATES

Carbohydrates are widely distributed in animal tissues, where they have been traditionally considered as structural components and reservoirs of nutrients. However, research in the last 15-20 years has also shown their involvement in diverse physiological processes[34]. A new discipline, glycobiology, has been established and devoted to the study of the role of carbohydrates in biological events. The rapid progress made underlines the importance of sugars in living systems, pushing them to the forefront of biological sciences, a position equal to proteins and nucleic acids. Nowadays, the general opinion is that there is no single function of oligosaccharides and indicative is the title of a review published by Varki in 1993: *Biological roles of oligosaccharides: all of the theories are correct*[35].

Unlike proteins and nucleic acids, polysaccharides can be highly branched and their monomeric units can be connected to one another by many different types of linkage. Each monosaccharide can be differently substituted and covalently bound to another one at different positions in the molecule. Furthermore, two monosaccharides can be linked together in two anomeric forms, the α- and β-linkage. Laine[36] reported that while the number of different hexapeptides that can be obtained connecting six different amino acids is 4.66×10^4, up to 1.95×10^{12} different hexasaccharides can be

formed binding six different monosaccharides. Therefore, theoretically, carbohydrates can carry per unit of weight much more information than nucleic acids and proteins. Moreover, the size of a sugar residue is larger than that of an amino acid. Thus, for instance, a relatively small oligosaccharide can shield large surface areas of the glycoprotein of which it is part. This does not only mean that the sugar moiety can strongly influence the protein properties but it also implies that cell surface oligosaccharides are among the first to be encountered by nearby molecules[37-39]. It is in these structural features that the enormous potential of carbohydrates to encode biological information, and therefore to act efficiently as recognition markers, residues[40,41]. When considering that glycoproteins are major components of the outer surface of mammalian cells and that they are essential for many biological processes, such as fertilisation, immune defence, viral replication and infection, cell growth and cell-cell adhesion[42], this hypothesis becomes even more plausible. Notable experimental evidence of the activity of sugars as recognition determinants has been collected throughout the years. Classical examples are the origin of the anticoagulant activity of heparin by the binding of a unique pentasaccharide to antithrombin III[43], or the results of the early studies of the Australian school[44], which showed that the presence of sialic acid residues on the surface of erythrocytes is an absolute requirement for the attachment of influenza virus; prior treatment of the red blood cells with neuraminidase abolishes the binding. It is noteworthy that the type of cell has a major role in determining the extent and type of protein glycosylation, which is both species and tissue specific[45]. Many enzymatic reactions are involved in the glycan-processing pathways[46]. This explains why the glycosylation pattern is influenced by cellular physiological changes due, for instance, to the occurrence of a disease (which may affect one or more enzymes in the cell). Thus, the altered

glycoform population of a given glycoprotein may be diagnostic of the disease responsible for the alteration itself. Abnormal glycosylation has been detected in cases of rheumatoid arthritis and cancer[47].

In principle, the same oligosaccharide can be attached to quite different proteins. However, the orientation of the oligosaccharide can significantly influence the properties of the glycoprotein. Therefore, the binding of the same oligosaccharide structure to different proteins can give rise to glycoproteins with markedly different characteristics, which again underlines that there is no single unifying function for carbohydrates[42].

1.3 LECTIN-CARBOHYDRATE INTERACTIONS

The idea that lectins and carbohydrates are excellent as cell recognition markers originates from the findings that both classes of compounds are commonly present on the cell surface and that sugars possess tremendous coding capacity. The ability of lectins to distinguish between subtle variations of oligosaccharide structure makes them perfectly suitable as decoders for the carbohydrate-encoded information. In other words, whilst sugars are able to carry the biological information, lectins are capable of deciphering the code. In this sense, the concept of lectin-ligand interaction can be seen as an extension of the lock-and-key hypothesis introduced by Fischer at the end of the nineteenth century to explain the specificity of the association of enzymes with their substrates.

Considerable experimental evidence is nowadays supporting the assumption that protein-carbohydrate interactions constitute the first step of a myriad of physiological events. So, for instance, old erythrocytes are cleared from the blood through the recognition of the galactosyl residues on their surface by the hepatocytes'

receptors (asialoglycoprotein receptors)[48]. With age the red blood cells become progressively desialylated. When the density of galactose moieties on the surface is high enough, the cell is removed from the circulation and destroyed (Figure 5). Galactose was first established to be the recognition marker for the mammalian hepatic carbohydrate receptor in 1976[49].

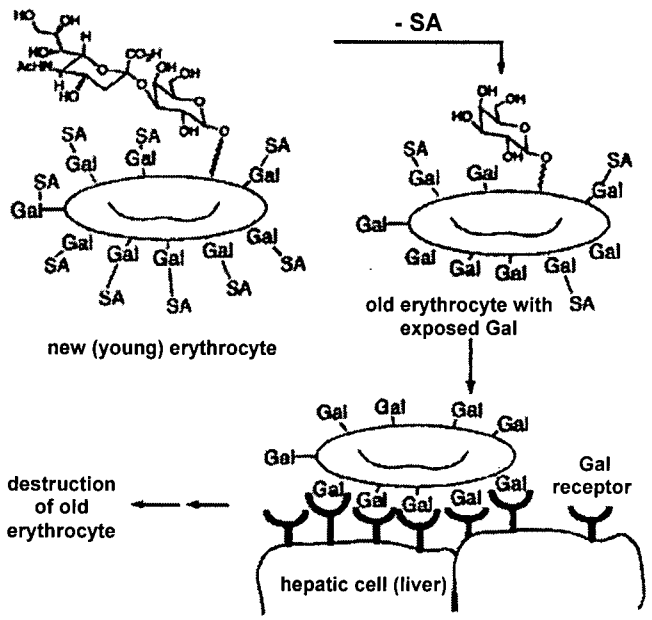


Figure 5 Clearance of non-sialylated erythrocytes from the blood by the liver.

The mammalian asialoglycoprotein receptor of hepatocytes is the prototype for the C-lectins, a large family of multidomain Ca^{2+}-dependent animal proteins. Members of this group are important components of the innate immune system, including the mannose-binding proteins, MBPs. The role of MBPs in host defence against viral pathogens has been demonstrated by several studies[50,51]. Like all other collectins – a group of soluble C-type lectins –, MBPs have a N-terminal collagenous domain and a globular C-terminal portion, containing the carbohydrate-recognition domain (CRD)[52]. Thus, while the C-terminal domain mediates the binding to various microorganisms, the N-terminal one interacts with the cell, triggering the

carbohydrate-dependent complement activation, which then results either in the killing of the pathogen[53,54] or in the deposition of complement components on the surface of the microorganism (opsonisation) which is then cleared by phagocytic cells like macrophages and neutrophils. MBPs have been shown to prevent human immunodeficiency virus (HIV) infection of H9 cells, probably by binding to oligomannose units of the envelope gp120 of the virus[55]. Unlike antibodies, whose reaction takes at least one to three days, these collectins react to infections or injuries immediately. A deficiency in MBPs has been claimed to account for a significant number of immunodeficiency cases in children, where the immune system is still not fully capable of mounting an efficient response[56].

Macrophages also express cell-surface galactose/*N*-acetylgalactosamine-specific C-type lectins (MMGL, mouse macrophage Gal/GalNAc-specific C-type lectins). These proteins have been shown to be involved in the recognition of cancerous cells by macrophages. Expression of MMGL, therefore, allows macrophages to be tumouricidal[57].

Selectins, another group of C-type lectins so called because they mediate selective contacts between cells, play a key role in the recruitment of leukocytes to inflammatory sites. During the inflammatory response, damage in tissues surrounding a blood vessel causes the influx of signalling molecules. Selectins are thus rapidly expressed on the inner surface of blood vessels (E-selectins) and on platelets (P-selectins). High specific binding of these proteins to sugars on white blood cell surfaces causes the leukocytes to adhere to the vessel walls, which, at this stage, express L-selectins. By interaction with L-selectins, leukocytes can now roll toward the site of damage, where they pass through to the surrounding tissue[58]. Whilst this is

a well-controlled process in healthy individuals, in excess it is a cause of arthritis, asthma and myocardial infarction.

Many microorganisms produce lectins, also referred to as adhesins, that are directly involved in the initiation of infections by mediating the adhesion to host cells through protein-carbohydrate interactions[59]. Intestinal epithelial cells of young piglets, which are susceptible to infection by *E. coli K99*, contain the glycolipid NeuGc-GM_3. Adult pigs, which are resistant to *E. coli K99*, lack this glycolipid[60].

1.3.1 MULTIVALENCY

The most striking features of lectin-monosaccharide interactions are that they are relatively weak, with association constants usually in the millimolar range for monosaccharides[61], and that they may show relaxed specificity, when compared to the strict nature of enzyme-substrate associations. The reason for this weakness lies in the solvent-exposed nature of the lectin binding-sites, which are shallow pockets making few direct contacts with the ligands[62]. There is, in fact, a significant difference in affinity between these shallow sites and deep sites, as is well illustrated by the influenza hamagglutinin lectin, which binds sialic acids with an approximately 1000-fold lower affinity than a neuraminidase found in the same virus[63]. Nevertheless, lectins exhibit both high affinity and exquisite specificity for oligosaccharide structures of glycoproteins and glycolipids on the cell surface. If this was not the case, lectins could not act as recognition molecules in biological processes. It has therefore been suggested that multiple protein-carbohydrate interactions are involved in the recognition event, giving the required high affinity and specificity[61]. Thus polyvalent associations occur throughout biology, showing a number of characteristics that monovalent interactions do not exhibit.

A well-studied polyvalent interaction involved in cell migration that may be important in metastasis is that between the galactosyl-transferase (GalTase) protein on one cell and the GlcNAc-terminated glycoproteins and glycolipids on another cell. Cells express different densities of both GlcNAc and GalTase, also depending on the time and the environment. The strength of cell-cell adhesion is related to the levels of GlcNAc and GalTase. It has been shown that a malignant cell expressing GalTase on its surface migrates along a concentration gradient of immobilised GlcNAc, whereas the same cell not expressing GalTase does not[64] (Figure 6).

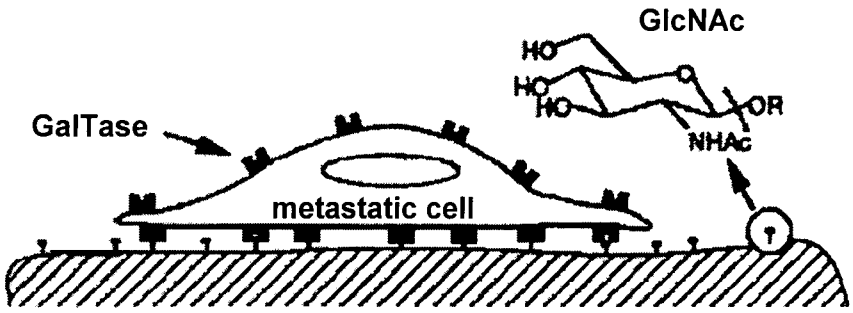


Figure 6 Adhesion of a metastatic cell to laminin. GalTAse is expressed on the edges of migrating cells and mediates cell migration by binding to oligosaccharides terminated with *N*-acetylglucosamine (GlcNAc).

In most cases, biological systems seem to use polyvalent interactions rather than a very strong single one. Therefore, multivalency is employed in Nature not only to achieve the necessary high affinity, but also to ensure the correct functioning of the cells.

It has been previously mentioned that old erythrocytes are removed from the blood through the association of the exposed galactose residues to the asialoglycoprotein receptors on hepatocytes. It is worth noting that only when the density of galactose moieties reaches a certain level the interaction is strong enough to lead to the erythrocyte clearance. Thus, in this case, polyvalency is used as part of a

timer that enables the body to judge the age of erythrocytes and select those that are old enough to be removed and destroyed.

All classes of antibodies have multiple equivalent receptor sites. Multivalency leads to high affinity binding to surfaces that show repeated epitopes, such as almost all invading pathogens. Mannose residues on the tail (the F_c portion) of the antibody interact with mannose receptors (the F_c receptors) on the surface of a macrophage (Figure 7a).

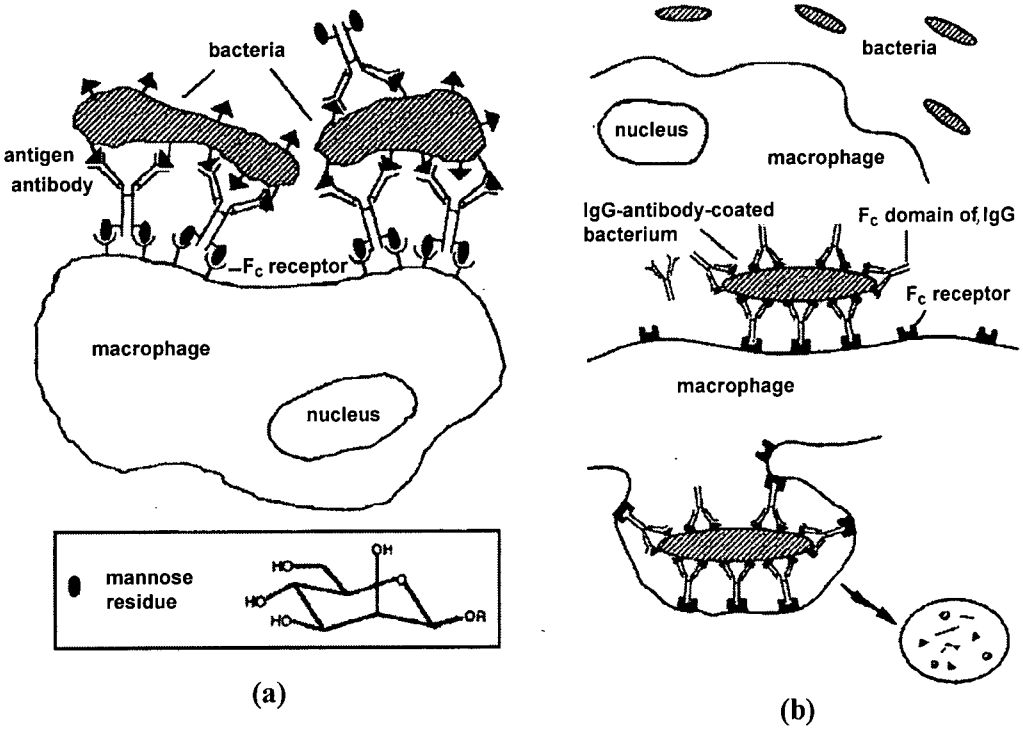


Figure 7 a), Polyvalent binding of a macrophage to a bacterium (or other pathogens) through antibodies, leading to a three-layered structure stabilised at both interfaces by polyvalent interactions; b), clearance of pathogens by antibody-mediated attachment to macrophages. The multivalent binding of the antibody-coated bacterium (or virus) to multiple F_c receptors of a macrophage activates the phagocytic process, leading to the degradation of the microorganism.

The interaction of a single F_c portion with its receptor seems to be too weak to induce a response by the macrophage, whereas multiple antibodies bound to the surface of

the bacterium can cross-link the F_c receptors, triggering an internal signal in the macrophage to ingest the infective agent[65] (Figure 7b).

In the receptor-mediated endocytosis of low-density lipoproteins (LDL), the ligand binds to a LDL receptor on the membrane of the cells. The resulting complex can then bind to an adaptor molecule in the cytoplasm, which can induce the formation of the clathrin-coated pit by multivalent association of clathrin and other coat-associated proteins. In this case polyvalency leads to the required clustering of receptors in the clathrin-coated pit[66] (Figure 8).

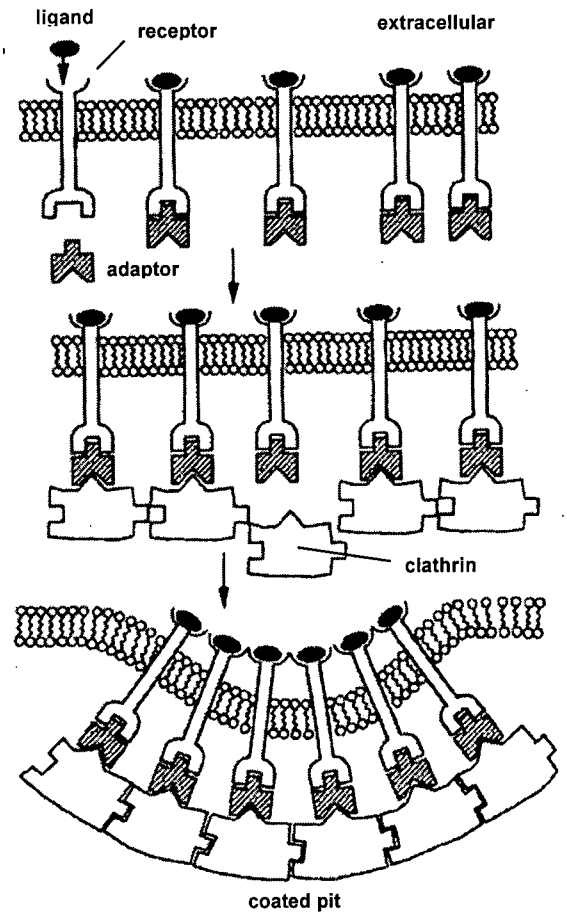


Figure 8 Induction of a specific geometric shape and molecular distribution: clustering of receptors in a clathrin-coated pit.

In some polyvalent systems, the interaction induces significant conformational changes that can represent biological signals. IgM is a class of antibodies with five bivalent structures extending randomly from a central ring. On binding decavalently to the bacterial surface, portions of the ring become more exposed to binding and differently organised. These portions initiate the chain of reactions that constitute the complement cascade by interaction with the pentavalent protein C1 in the blood. This cascade results in the death of the microorganism. Binding of IgM to monovalent ligands does not lead to any conformational change and, therefore, does not induce the complement cascade. In this process, polyvalency is required not for tight binding, but for signalling and activation through a conformational change[66] (Figure 9).

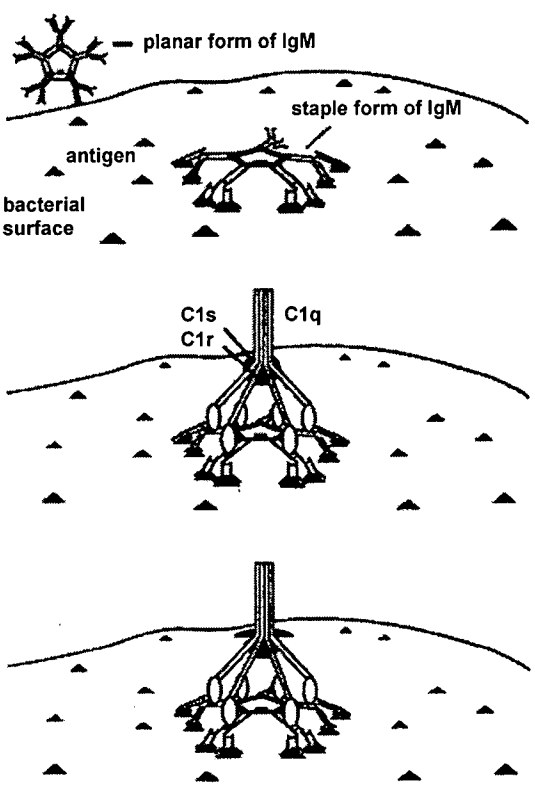


Figure 9 Conformational change of a pentameric immunoglobulin IgM upon multivalent binding to antigens on the surface of the bacterium. The change activates the complement cascade.

The interaction between an allergen and a mast cell proceeds through cross-linking of the receptors. The mast cells present multiple copies of bivalent IgE and can degranulate, *i.e.* release histamine that promotes local inflammation, in response to a wide range of allergens. Monovalent ligands for IgE do not trigger degranulation, whereas cross-linked suface-bound IgE bind to an intracellular protein, initiating a signal cascade that ends in degranulation[67] (Figure 10).

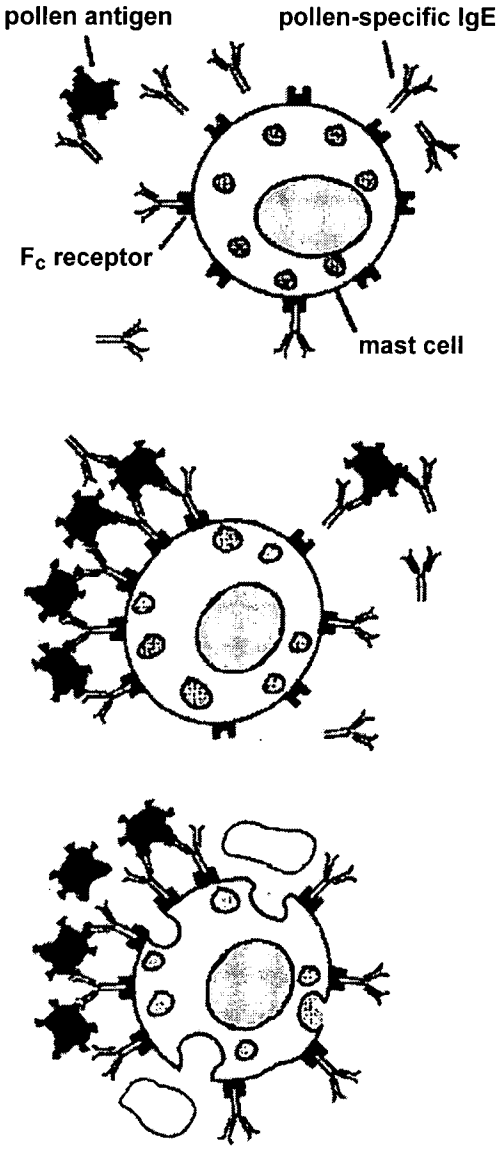


Figure 10 Degranulation of mast cells after binding of a polyvalent allergen.

1.3.2 ENERGETICS

A full understanding of the mechanism of protein-carbohydrate association at a molecular level requires the elucidation of both structural and energetic aspects of the process. As briefly discussed in the previous sections, a large amount of structural data is now available. However, the link between structure and energetic properties remains too obscure to allow a description of the energetics of the interaction from structural information[25]. Historically, the strength of protein-ligand complexation has been determined by hemagglutination and precipitation inhibition assays, although several spectroscopic techniques, affinity chromatography and equilibrium dialysis have been also utilised. The so obtained values of the binding constant, K, and of the free energy of binding, ΔG, give little information about the forces involved in the complex formation. Moreover, the enthalpy and entropy of association are derived only indirectly from van't Hoff analyses of the temperature dependence of the free energy:

$$\frac{d \ln K}{dT} = \frac{\Delta H}{RT^2} \qquad (1)$$

The employment of this equation requires the assumption that the enthalpy of binding, ΔH, is not a function of the temperature, *i.e.* $\Delta C_P = 0$, being $\left(\frac{\partial H}{\partial T}\right)_P = \Delta C_P$. However, the heat-capacity change, ΔC_P, for ligand binding in aqueous media is non-zero, making doubtful the accuracy of the calculated values.

The accessibility of commercial high-resolution microcalorimeters has provided a direct method of evaluation of accurate thermodynamic data that can complement the structural information. Isothermal titration microcalorimetry, ITM, is

the only technique that allows direct determination of K and ΔH in a single experiment. Values of ΔC_P are also obtained performing titrations at different temperatures. ΔC_P is a powerful indicator of solvent behaviour during binding[68] and provides invaluable information about the driving forces for association. Accurate values of ΔC_P are available only by microcalorimetry.

1.3.2.1 MONOVALENT LIGANDS

Calorimetrically determined thermodynamic constants for lectin- carbohydrate interactions are listed in Table 1. In almost all cases the enthalpy of binding is more negative than the free energy. It is noteworthy that this pattern is generally found for association in aqueous solution and it is not a special feature of lectin-ligand complexation[69,70]. It has already been pointed out that water molecules are strongly involved in protein-sugar interactions, giving rise to bridges between the saccharide hydroxyl groups and the protein amino acid residues. Watson *et al.*[71] have reported that in the case of complexation of glucose with glycogen phosporylase water-mediated hydrogen bonds are as strong as the direct ones, reinforcing the concept of structural water as an extension of the protein surface. However, the displacement of a large number of water molecules during binding can also be observed. The effect of this release of water has been long debated. Numerous models have been formulated in order to explain the thermodynamic properties of apolar organic molecules in water, particularly for the surprising cases where the most significant term opposing dissolution is enthalpic rather than entropic[25]. It has been proposed that although the strength of hydrogen bonds in the hydration shell of the protein is higher than in bulk solvent, the fraction of broken bonds is also higher, due to geometric constraints imposed on water by the solute. Thus, the return of water of solvation to bulk water

Table 1 Calorimetrically derived thermodynamic parameters of lectin-carbohydrate interaction.

Lectin	Carbohydrate	$10^3 \times K$ M^{-1}	ΔG kJ mol^{-1}	ΔH kJ mol^{-1}	ΔS J mol^{-1} K^{-1}	ΔC_P J mol^{-1} K^{-1}	Ref
WBAI [a]	Gal	1.2	-17.5	-24.3	-23		[72]
	GalNAc	7.2	-22.0	-28.0	-20		
	MeαGal	6.6	-21.8	-23.5	-6		
	MeβGal	1.0	-17.1	-19.7	-9		
Con A	MeαMan		-22.2	-27.6	-18.3	-251	[69]
	Manα1,6(Manα1,3) Man [b]		-30.1	-41.0	-36.5	-460	
Con A	MeαMan	11.9	-22.9	-29.2	-22		[73]
	Man	2.1	-18.6	-23.9	-18		
	MeαGlc	2.7	-19.2	-18.1	+4		
	Glc	0.6	-15.4	-17.1	-6		
pea	MeαMan	1.9	-18.3	-27.3	-31		
	Man	0.9	-16.5	-24.8	-29		
	MeαGlc	0.6	-15.5	-13.1	+8		
Con A	MeαMan	8.2		-34.3	-40.7		[74]
	Manα1,6(Manα1,3) Man	49.0		-60.3	-92.7		
DGL [c]	MeαMan		-20.5	-32.6	-40.7	-234	[75]
	Manα1,3Man		-26.4	-47.7	-71.6	-167	
	Manα1,6Man		-21.0	-36.0	-50.5	-92	
	Manα1,6(Manα1,3) Man		-34.3	-54.4	-67.4	-402	
Con A	MeαMan		-22.2	-28.5	-21.0	-209	
	Manα1,3Man		-25.2	-31.0	-19.6	-460	
	Manα1,6Man		-22.2	-28.9	-22.5	-184	
	Manα1,6(Manα1,3) Man		-31.0	-42.7	-39.3	-389	
SBA [d]	MeβGal	0.5		-44.4	-96.2	-393	[76]
	GalNAc	9		-39.7	-57.2		
	MeβGalNAc	22		-58.1	-111.6	-418	
	LacNAc	0.7		-34.3	-60.0		
ECorL [e]	MeβGal	0.4		-18.4	-11.2		
	GalNAc	1.2		-29.7	-40.4		
	MeβGalNAc	1.3		-28.5	-34.9		
	LacNAc	4.2		-45.6	-82.3	+393	
Gal-1 [f]	LacNAc	6.2		-27.6	-19.5	-377	
C2S-Gal-1	LacNAc	2.9		-11.7	+27.1		
N-Gal-1	LacNAc	8.7		-2.5	+66.5		
ECorL	Gal	1.6	-18.2	-13.7	+15.3		[77]
	MeαGal	1.4	-18.1	-21.6	-12.0		
	MeβGal	0.7	-16.3	-18.2	-6.6		
	GalNAc	1.3	-17.9	-23.0	-17.1		
	Lactose	1.9	-18.8	-41.2	-75.4		
	LacNAc	9.7	-22.7	-47.1	-83.2		
	2'-FL [g]	3.7	-20.3	-18.0	+7.7		
	MeαDNSGalN [h]	351.5	-31.7	-23.1	+30.0		
	Fucose	0.5	-15.2	-4.7	+35.2		
Gal-1 [i]	Galβ1,4Glc	6.4	-21.7	-20.9	0		[78]
	Galβ1,4Fruc	8.2	-22.4	-34.4	-40		
	Galβ1,4Man	11.3	-23.0	-35.2	-45		
	Galβ1,4Ara	5.4	-21.3	-36.8	-52		
	Galβ1,4GlcNAc	22.2	-24.7	-35.9	-38		
	Galβ1S1βGal	11.6	-23.1	-46.4	-78		
PNA [j]	Galβ1,3GalNAc	20.6	-24.7	-59.0	-115.5		[79]

[a] Basic lectin from winged bean. [b] 3,6-di-*O*-(α-D-mannopyranosyl)-α-D-mannopyranoside. [c] Lectin from *Dioclea grandiflora*. [d] Soybean agglutinin. [e] Lectin from *Erythrina corallodendron*. [f] From Chinese hamster ovary cells. [g] 2'-Fucosyllactose. [h] Methyl-α-*N*-dansylgalactosaminide. [i] From bovine spleen. [j] Peanut agglutinin.

might be enthalpically driven, providing a favourable contribution to the enthalpy of complexation[80]. Monte Carlo simulations, carried out to investigate the structure of water near the surface of the lectin binding site, have shown highly disordered water in the proximity of the protein surface. The phenomenon has been explained assuming that the protein surface could not be complementary to any low-energy structure that the water could adopt[81]. Chervenak and co-workers[68] have reported a calorimetric evaluation of the thermodynamic binding parameters for several systems in light and heavy water. In all cases desolvation was found to contribute to a significant fraction of the binding enthalpy: *ca.* 25% for lectin-sugar association. To date, the role of protein-ligand hydrogen bonding *versus* solvent reorganisation as contributors to the enthalpy of binding remains the most polemic area of discussion.

Another common feature of lectin-carbohydrate interactions is the strong linear enthalpy-entropy compensatory behaviour (Figure 11).

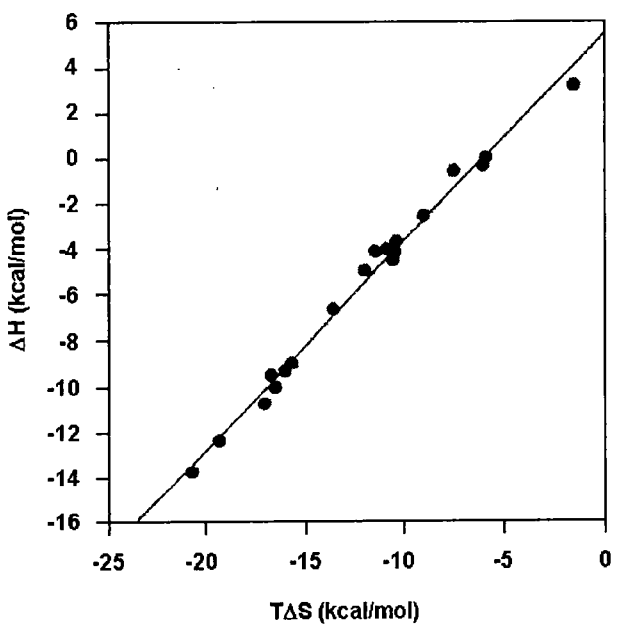


Figure 11 Enthalpy-entropy compensation behaviour for the interaction of WBAII with Fucα(1,3)Galβ(1,4)GlcNAc-OMe and a series of deoxy and -OMe derivatives. The slope of the plot is 0.90 and the correlation coefficient 0.98[82]. A slope value smaller than unity indicates that the binding is dominated by entropy.

This offset has been interpretated both in terms of changes in the degrees of freedom of the ligand upon binding[73,83] and in terms of solvent reorganisation[84,85]. Chervenak *et al.*[75] have determined the configurational and solvation entropies for the formation of a series of complexes with Con A and DGL (Table 2).

Table 2 Entropic contributions to the binding of mannose oligosaccharides to Con A and DGL.

System	ΔC_P J mol^{-1} K^{-1}	ΔS J mol^{-1} K^{-1}	ΔS_{solv} J mol^{-1} K^{-1}	ΔS_{config} J mol^{-1} K^{-1}
Con A				
MeαMan	-209	-20.5	+53.6	-40.6
Manα1,6Man-OMe	-184	-3.8	+47.3	-17.6
Manα1,3Man	-460	-20.1	+118.0	-104.6
Manα1,6(Manα1,3)Man	-389	-51.9	+118.0	-136.4
DGL				
MeαMan	-234	-41.1	+60.2	-67.8
Manα1,6Man	-92	-50.3	+23.4	-40.2
Manα1,3Man	-167	-72.0	+42.7	-81.2
Manα1,6(Manα1,3)Man	-402	-110.5	+102.9	-179.9

The total entropy of the process can be separated into terms for changes in solvation and losses of configurational, rotational and translational degrees of freedom.

$$\Delta S = \Delta S_{solv} + \Delta S_{config} + \Delta S_{rot} + \Delta S_{trans} \tag{2}$$

The sum of the first two terms has been referred to as the unitary contribution, while the sum of rotational and translational entropies has been defined as the cratic contribution[86]. Knowing that:

$$\left(\frac{\partial S}{\partial T}\right)_P = \frac{\Delta C_P}{T} \tag{3}$$

ΔS_{solv} can be expressed as follows:

$$\Delta S_{solv} = \Delta S^*_{solv} + \Delta C_P \ln\left(\frac{T}{T^*}\right) \tag{4}$$

where T^* is the temperature at which there is no solvent contribution and it is equal to 385.15 K. ΔS^*_{solv} contains protonation and electrostatic contributions and for lectin-ligand interaction can be considered equal to zero. Murphy *et al.*[87] have proposed that for the formation of a 1:1 complex with 1 M as the standard state, the entropic cratic contribution, $\Delta S_{rot} + \Delta S_{trans}$, is equal to -33.5 J mol^{-1} K^{-1}. Thus, accurate measurements of ΔC_P provide ΔS_{solv} and, consequently, ΔS_{config}. For the systems reported in Table 2, both solvation and configurational entropies roughly scale with the size of the ligand, as a result of both the number of water molecules released and the number of degrees of freedom restricted during binding. Moreover, unfavourable configurational entropies are responsible for the observed enthalpy-entropy compensation (shown in Table 1). The more favourable entropy of association of the 1→3 dimannoside to Con A than to DGL has been attributed to a greatly reduced solvation contribution in the latter case.

It is now clear that the heat-capacity change contains important information about the mechanism of protein-carbohydrate complexation. The factors affecting ΔC_P have been discussed extensively and, at present, the term is usually thought to reflect solvent reorganisation effects and, to a lesser extent, changes in the protein vibrational modes[88]. For lectin-sugar interactions, ΔC_P values are usually small (≤ 400 J mol^{-1} K^{-1}) and negative. The significantly larger heat-capacity changes accompanying antibody-saccharide associations compared to lectin-sugar complexations can be interpreted in terms of structural differences between the binding sites of the two classes of biological polymers[77]. The antibody combining site contains several aromatic amino acids and the burial of these hydrophobic residues represents the driving force for protein-antigen binding[89]. However, the entropy of binding is still negative, offsetting the enthalpic gain. Mutagenesis studies have revealed that part of

the entropic loss may be associated with the reduced mobility of amino acid side chains in the antibody combining site[90]. In this case, therefore, the enthalpy-entropy compensation is also a function of protein structure. Considering again the data reported in Tables 1 and 2[75], considerably different mechanisms of association can be observed for the binding of the same oligosaccharide to the two lectins. Small but significant differences in the primary sequences of the two proteins may affect the flexibility of the loops forming the two binding sites. Although the lectin binding sites are greatly preorganised, several amino acid side chains move considerably to achieve the optimal hydrogen bonds and van der Waals contacts with the ligand. Thus, an enhanced flexibility would favour hydrogen bonding, but, at the same time, a more severe entropic loss would be observed. However, since the values of ΔC_P are small, little changes in protein conformation are usually involved, as confirmed by X-ray crystallographic determinations.

The larger $-\Delta H$ values obtained for the binding of 3,6-di-*O*-(α-D-mannopyranosyl)-α-D-mannopyranoside to Con A and DGL compared to methyl-α-D-mannopyranoside indicate the presence of extended binding sites. The trimannoside is, in fact, the minimum carbohydrate epitope that completely fills the combining pocket, giving rise to the maximum number of hydrogen bonds and van der Waals interactions[91]. This oligosaccharide represents the core structure of all *N*-linked glycoproteins[75].

Despite the general feature of being enthalpically driven with an unfavourable entropic contribution, few lectin-carbohydrate systems show positive entropy changes. As reported in Table 1, most of the carbohydrates that bind to ECorL show negative ΔS, but positive entropic changes are observed for galactose, fucose, methyl-α-*N*-dansylgalactosaminide and 2′-fucosyllactose[77]. Favourable entropies are

associated with non-polar contacts, for example between the aromatic dansyl moiety at C-2 of MeαDNSGalN and Trp 135 in the binding site of the lectin. From X-ray crystallographic studies[92], the increase of entropy associated with the binding of galactose is explained by the release of tightly bound water molecules, not still fully compensated by opposing entropically disadvantaged factors.

Differences in the thermodynamics of binding for homologous lectins, for which most of the amino acid residues in the binding sites are conserved, might indicate the indirect but critical role of non-conserved residues away from the combining pocket[93]. Effects of single-site mutations on the conformation of lectins have been investigated by Siebert and co-workers[94]. The mutation of sites far removed from the carbohydrate binding site in Galectin-1 from CHO cells strongly affects its thermodynamics of complexation with LacNAc[76]. As shown in Table 1, the two mutants C2S-Gal-1 and *N*-Gal-1, exhibit significantly lower -ΔH values than Gal-1, whereas the entropic contributions become favourable.

Importantly, studies carried out with monodeoxy analogues of 3,6-di-*O*-(α-D-mannopyranosyl)-α-D-mannopyranoside have shown that K and -ΔH for the analogues are lower than for Manα1,6(Manα1,3)Man[95]. The sum of the $\Delta\Delta H$ and $\Delta\Delta G$ values for the hydroxyl groups of the trimannoside, obtained from the monodeoxy analogues, do not correspond to the measured ΔH and ΔG for this ligand. This means that the magnitude of $\Delta\Delta H$ and $\Delta\Delta G$ represents not only the loss of the hydrogen bond involved, but also differences in the solvent and protein contributions to the binding of Manα1,6(Manα1,3)Man and of the deoxy analogues.

No correlation was found between the calorimetrically derived thermodynamic parameters of a series of ligands for the bovine spleen Galectin-1 and the number of protein-sugar close contacts[78]. Interestingly, binding enthalpies calculated from

changes in the solvent-accessible areas of the Galectin-1 binding site upon complexation showed poor agreement with the calorimetric values, again reflecting the importance of factors other than the burial of hydrophobic surface area.

Everything mentioned above shows the complex balance of forces before and after the binding event. It is therefore clear why both detailed structural and energetic information is essential in order to design ligand mimics. Attempts to overcome certain thermodynamic barriers through the planning of the glycoconjugate structure may not always result in the expected high-affinity association, due to the offset of different contributions[96].

1.3.2.2 MULTIVALENT LIGANDS

The "cluster glycoside effect", defined as "the enhancement in the activity of a multivalent ligand beyond what would be expected due to the increase in sugar local concentration (statistical effect) alone"[97] is nowadays generally accepted. Despite the numerous observations of the phenomenon, a molecular interpretation of the effect is difficult and, at present, its physical origin is still not well understood. Efforts to provide a molecular basis for the cluster glycoside effect have been complicated by two issues: 1) polyvalent ligands are often polydisperse and structurally ill-defined and 2) methodologies used to evaluate protein-carbohydrate binding measure several phenomena, including protein-carbohydrate association. Moreover, there may exist a relationship between the measured magnitude of the cluster glycoside effect and the assay utilised for the determination[98]. In a first approximation, a trend of increased enhancement with increasing valency can be drawn.

In principle, at least two models of association can be described: an intramolecular, or chelate, binding and an intermolecular aggregative process. In the

former a multivalent ligand spans a number of binding sites on a single protein, while in the latter the spanned binding sites belong to different receptor molecules, resulting in aggregates that may or may not precipitate.

Considering, for simplicity, a bivalent ligand, the thermodynamic parameters describing a chelate complexation can be expressed as follows[99]:

$$\Delta J_{bi} = 2\Delta J_{mono} + \Delta J_i \tag{5}$$

where: ΔJ_{bi} = any thermodynamic parameter for bivalent complexation

ΔJ_{mono} = corresponding term for monovalent association

ΔJ_i = interaction term

Interaction energies have been traditionally considered in entropic terms, where translational and rotational savings and conformational penalties have to be taken into account. As the translational and rotational entropies of a molecule are, respectively directly and indirectly, proportional to the logarithm of the molecular weight, the binding of two (or more) ligands produces a multivalent ligand with translational and rotational entropy roughly equivalent to that of the monovalent ligand. Therefore, the binding of a bivalent ligand is characterised by a favourable contribution to the entropic term equal to the translational and rotational entropy of the corresponding monovalent ligand. This term presumably accounts for a large part of the interaction free energy. On the contrary, the loss of conformational degrees of freedom of the ligand upon binding results in an unfavourable contribution to the overall entropy. Enthalpy also plays an important role in determining the overall affinity of a multivalent ligand. Alteration of the ligand position within the binding site would result in an unfavourable contribution, while the contribution of the linker may be

favourable due to its favourable interactions with the protein surface (at the periphery of the binding site or over the space separating the spanned binding sites) or unfavourable as a result of disadvantaged steric interaction. This latter aspect can result in a significant unfavourable contribution, due to the lock of linker dihedrals into gauche or eclipsed orientations, which can, in turn, preclude the complexation by an intramolecular mode. Considering what has just been said and combining it with the fact that distances between binding sites on the receptor molecule are generally of the order of tens of Angstroms (65 Å in the canonical dimer of Concanavalin A), it appears clear that the achievement of chelate complexation is challenging.

Most of the multivalent ligands reported so far reach the observed affinity enhancement by means of an aggregative process, where the formed cross-linked complexes may be stabilised by a range of forces, including protein-protein interactions. Alternatively, a diminished solubility of the aggregate and its precipitation would lead to an apparently enhanced affinity through a coupled equilibrium[100].

Few calorimetric studies of multivalent carbohydrate-protein interactions (indeed few for any multivalent ligand) have been reported so far[100-106]. Moreover, ITM has been limited to small or dendritic glycoconjugates possessing a maximum number of carbohydrate residues equal to six. Brewer *et al.*[103] have recently reported the determination of the thermodynamics of binding of small multivalent ligands (with structural valency up to 4) to Con A and DGL. The structures of those that showed an affinity enhancement relative to the corresponding monovalent ligands are reported in Figure 12. The thermodynamic parameters determined using Con A as receptor protein are reported in Table 3 (blue), together with those reported by Toone

et al.[102] (red) for the interaction of the two series of ligands (dendritic compounds containing up to six binding epitopes) shown in Figure 13 with the same lectin.

Table 3 Thermodynamic binding parameters of Con A with multivalent ligands.

Compound	$10^4 \times K$ M^{-1}	ΔG kJ mol^{-1}	ΔH kJ mol^{-1}	$T\Delta S$ kJ mol^{-1}	n*
Manα1,6(Manα1,3)Man	39	-31.8	-61.5	-29.7	1.0
1	286	-36.8	-96.6	-59.8	0.53
2	250	-36.4	-109.6	-73.2	0.53
3	420	-37.6	-121.3	-83.7	0.51
4	1350	-40.6	-221.8	-181.2	0.26
MeαMan	0.76	-22.2	-27.6	-5.4	1
5	0.92	-22.6	-31.0	-8.4	1
6	0.80	-22.2	-31.4	-9.2	1
7	0.79	-22.1	-32.6	-10.5	1
8	0.75	-22.2	-17.6	+4.6	1
9	ND				
10	0.80	-22.2	-26.8	-4.6	1
11	0.86	-22.6	-32.2	-9.6	1
12	0.47	-20.9	-29.7	-8.8	1
13	6.2	-27.6	-9.6	+18.0	1
14	150.0	-35.5	-5.4	+30.1	1

* Blue: [Ligand] = molar concentration of the overall ligand molecule; Red: [Ligand] = valency-corrected concentration.

As previously discussed, the interaction of Con A with monovalent carbohydrates was found to be enthalpically driven, with an unfavourable entropic contribution[69]. Despite the different behaviour of the thermodynamic parameters going from the mono- to the multivalent ligands in each series, the affinity enhancement observed for ligands **1-4** and **13-14** was in both cases explained by means of an intermolecular aggregative mechanism. Assuming non-cooperativity of the different epitopes in the multivalent carbohydrates, Toone *et al.* explained the thermodynamics of lectin-ligand association (ligands **13-14**) hypothesising an initial exothermic protein-carbohydrate interaction, which proceeds with thermodynamic parameters equivalent to those of the corresponding monovalent saccharide, coupled with or followed by an endothermic, entropically driven, aggregation process. The diminution of $-\Delta H$ was therefore considered the thermodynamic signature for the intermolecular aggregation process[98,102].

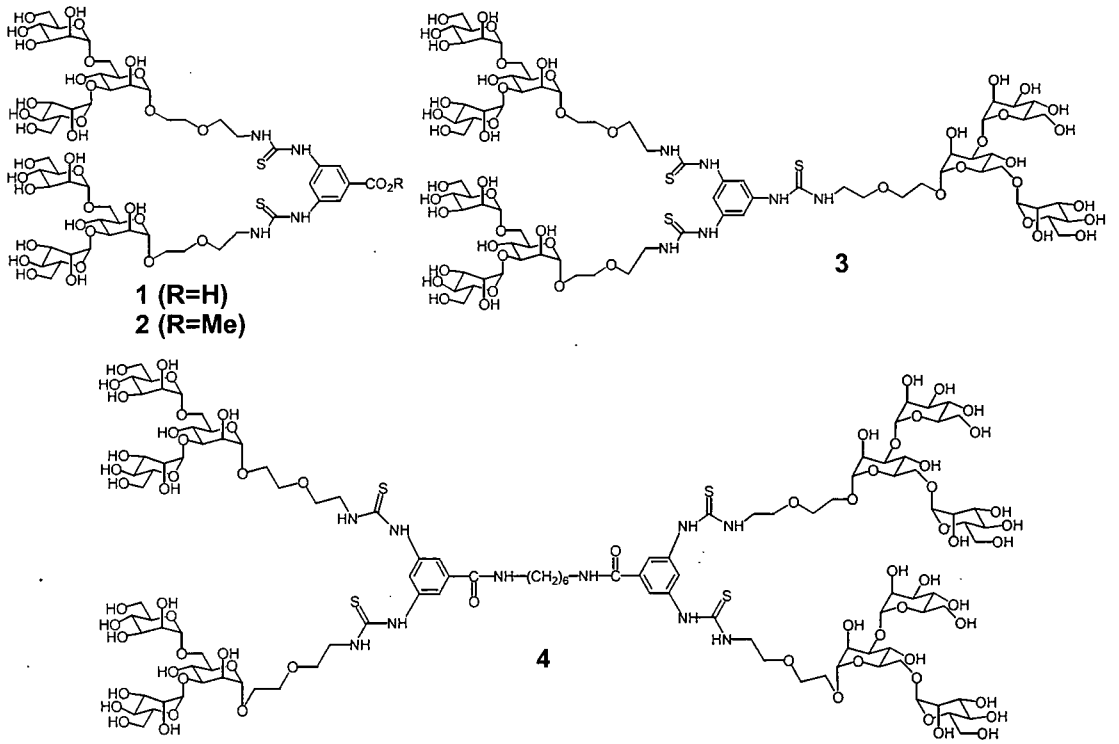


Figure 12 Multivalent ligands used by Brewer *et al.*[106]. The corresponding thermodynamic binding parameters are reported in blue in Table 3.

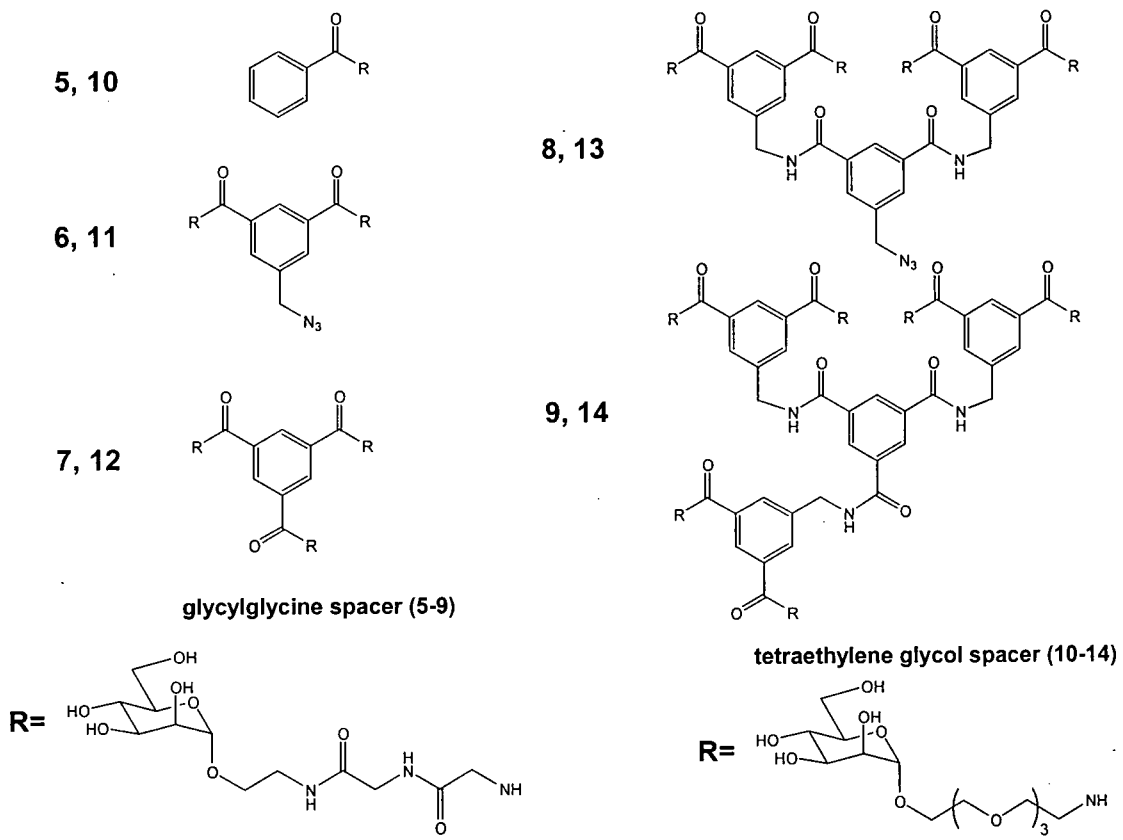


Figure 13 Dendritic ligands used by Toone *et al.*[102]. The corresponding thermodynamic binding parameters are reported in red in Table 3.

Studies previously carried out by the Toone and co-workers employing another series of dendritic multivalent ligands[101], where crystals of both bi- and trivalent ligands were grown in the presence of succinylated Con A, showed that the carbohydrate orientation within the binding site was identical to that of the corresponding monovalent ligand. Interestingly, ligands with valencies greater than three showed affinity enhancements in agglutination assay but not by ITM. On the other hand, tetra- and hexavalent carbohydrates showed substantially diminished enthalpy (in absolute value) and visible cloudiness. Therefore, the IC_{50} values were found to correlate with the calorimetrically derived entropies, but not with the free energies of ligand binding. The authors deduced that the agglutination assay did not evaluate the strength of protein-carbohydrate interaction, but rather the ability of multivalent ligands to drive aggregation processes. Based on these results, the overall thermodynamic parameters could be expressed as the sum of two contributions, the actual protein-sugar interaction – characterised by thermodynamic values equal to those determined for the corresponding monovalent ligand – and the aggregation process. Enthalpies of aggregation were thus determined.

Despite the common explanation for the observed enhancement of the binding affinity through an intermolecular aggregative mechanism, the results reported by Brewer *et al.* are rather different from those of Toone and co-workers. In the former case, an increase of $-\Delta H$ was observed going from the monovalent to the multivalent ligand (**1**-**4**). With the exception of compound **3** (which was found to behave as a divalent ligand despite its structural valency being equal to three[103,105]), ΔH scaled proportionally to the number of epitopes in the multivalent ligand, whereas ΔS did not. The entropy variation was instead more negative than the sum of ΔS of individual epitopes. According to the authors, this behaviour of ΔS indicated the binding of the

multivalent ligands to separate receptor molecules (aggregation). Moreover, the non linearity of Scatchard and Hill plots was explained by a negative cooperativity for the binding of multivalent ligands[104], which was experimentally demonstrated by measuring the microscopic thermodynamic parameters of each carbohydrate residue in the ligand by means of reverse isothermal titration microcalorimetry[105]†. In this case solutions of the mono-, bi- and trivalent ligands were placed into the cell and Con A used as the titrant. For instance, the binding constant, K_1, of the first epitope of the bivalent ligand was found to be one order of magnitude greater than those of the second epitope and of the corresponding monovalent carbohydrate. The average of the two microscopic binding constants determined agreed with the macroscopic binding constant determined using the sugars as injectants. The enthalpic contribution of the two epitopes was essentially the same and very close to that for the monovalent ligand. Therefore, the enhancement in affinity observed was associated with a favourable entropy of binding of the first epitope relatively to the second one. However, while the sum of the two enthalpic contribution was very close to the enthalpy determined by "normal" microcalorimetry, the sum of the two microscopic entropies was much lower, in absolute value, than the one reported when the ligand was injected into the lectin solution. Hemagglutination inhibition measurements perfectly correlated with the results obtained by ITM. The authors provided some hypotheses of explanation for the differences between their results and those reported by Toone and co-workers. For instance, they underlined the difference in the way adopted to express the concentration of the ligands. While Toone *et al.* considered the epitope equivalents, Brewer *et al.* referred to the molar concentration of the ligand molecule. Their arguments against the use of valency-corrected concentrations were

† For general aspects of isothermal titration microcalorimetry, see Chapter 4, Section 4.1.1.

the observed negative cooperativity of binding and the determined difference between the structural and functional valency[104].

Most examples of intramolecular/chelate binding involve either bacterial toxins or polymeric ligands. Studies have shown that the enhancement in binding affinity is much higher when different carbohydrate residues of the ligand bind to sites on the same receptor molecule.

In 1983, Lee *et al.*[107] reported that a synthetic tetraantennary undecasaccharide showed an inhibition constant for the hepatic Gal/GalNAc receptor 10^6 times greater than an equivalent monoantennary trisaccharide, despite the only 4-fold statistical increase in absolute galactose concentration. It was therefore pointed out that the number of Gal residues per cluster and their branching mode were major determinants of the binding affinity of a ligand to the mammalian hepatic lectin.

An extraordinary affinity enhancement achieved through multivalency was reported by Rao *et al.*[108] in 1998. The binding constant of a trivalent derivative of vancomycin, $R_t\mathbf{V}_3$, to a trivalent derivative of D-Ala-D-Ala (DADA), $R'_t\mathbf{L}'_3$, was *ca.* 10^{16} M^{-1} compared to the value of *ca.* 10^6 M^{-1} of the corresponding monovalent analogues (Figure 14a). Calorimetric studies showed that both enthalpic and entropic terms were approximately three times those determined for the interaction of $R_t\mathbf{V}_3$ and the monovalent ligand $R'_t\mathbf{L}'$ (Table 4). In a subsequent study[109], the same authors reported that the efficiency of a dimeric analogue of vancomycin, $\mathbf{V}$-R_d-$\mathbf{V}$, in inhibiting the growth of vancomycin resistant enterococci was 100 times greater than that of monomeric vancomycin, **V**. Moreover, the binding constant of $\mathbf{V}$-R_d-$\mathbf{V}$ to Lac-R'_d-Lac, whose structures are reported in Figure 14b, was approximately 40 times tighter than that of **V** to Ac_2KDADLac (analogous monovalent system). The corresponding thermodynamic binding parameters are reported in Table 4.

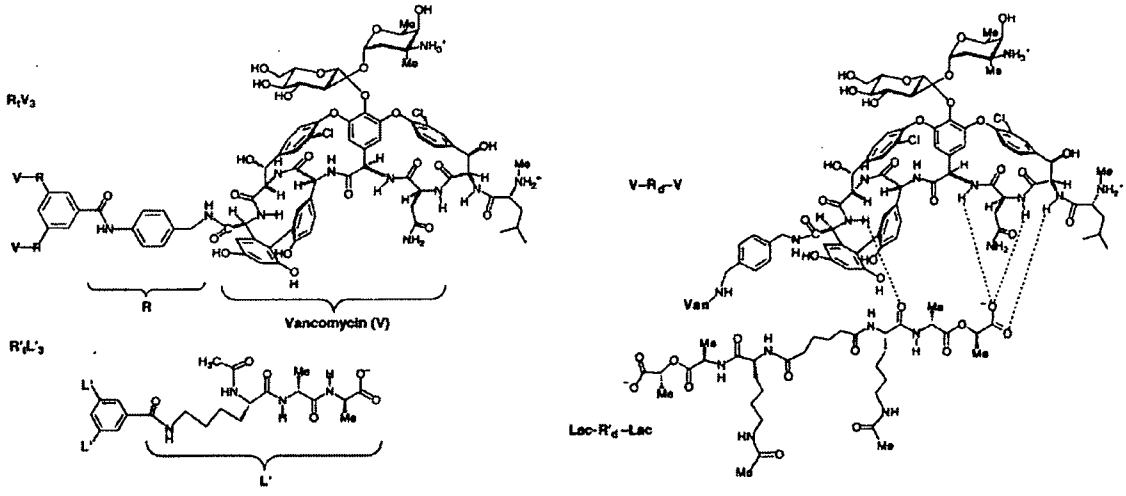


Figure 14 a), Structures of the trivalent derivatives of vancomycin, R_tV_3 and $R'_tL'_3$; b), structures of the dimeric vancomycin, **V**-R_d-**V** and dimeric lactate ligand, Lac-R'_d-Lac.

Table 4 Thermodynamic parameters of binding of vancomycin derivatives to mono- and multivalent ligands.

Receptor	Ligand	$10^6 \times K$ M^{-1}	ΔG kJ mol^{-1}	ΔH kJ mol^{-1}	$T\Delta S$ kJ mol^{-1}
R_tV_3	$R'_tL'_3$	2.5×10^{10}	-94.1	-166.9	-72.8
R_tV_3	R'_tL'	9.1×10^{-1}	-33.9	-51.9	-18.0
R_tV_3	**L**	3.7×10^{-1}	-31.8	-50.2	-18.4
Vancomycin	$R'_tL'_3$	2.9	-36.8	-73.2	-36.4
Vancomycin	**L**	6.3×10^{-1}	-33.0	-50.1	-17.1
V-R_d-**V**	Lac-R'_d-Lac	2.3×10^{-2}	-25.1	-27.2	-2.1
Vancomycin	Ac_2KDADLac	5.9×10^{-4}	-15.9	-15.5	+0.4

Most of the currently synthesised polymeric and dendritic multivalent ligands show a "random" multivalency, which rarely allows an affinity enhancement higher than 1,000-fold. Moreover, as already pointed out, the reasons underlying this activity gain are poorly understood[66]. In a different approach, structural information about the spatial arrangement of the target binding sites are taken into account for the design of ligands ideal for maximising the interaction with the receptor.

The modular synthesis of multivalent ligands of the heat-labile enterotoxin (LT) from *E. coli* was recently reported by Fan and co-workers[110]. The five B

subunits of the toxin present a 5-fold symmetric configuration. In this study, pentavalent ligands constituted by a semirigid "core" which can adopt a conformation close to 5-fold symmetry, flexible "linkers" that project toward the receptor binding sites and "fingers" represented by 1-β-amidated D-galactose, were synthesised. Different lengths of the linker were tested and the receptor-ligand interaction analysed by an enzyme-linked immunosorbent assay, ELISA. All pentavalent ligands led to significant affinity gains compared to the corresponding monovalent ones. The best one showed an IC_{50} 10^5-fold higher than galactose, approaching the affinity of the oligosaccharide portion of the natural ganglioside GM1. The pentavalent ligand also showed a 10^4-fold gain compared to the corresponding monovalent ligand, or 2,000-fold on a valency-corrected basis. Dynamic light scattering ruled out the possibility of an aggregative process, supporting the expected formation of a 1:1 toxin-ligand complex. The analysis of the effect of linker-length on the affinity showed that the greatest enhancement was detected for the ligand whose linker effective length[111] best matched the distance between nonadjacent binding sites.

One of the most striking examples of affinity enhancement was recently reported by Kitov *et al.*[112]. Shiga-like toxins SLT-I and SLT-II were inhibited by a decavalent ligand, designated STARFISH, whose structure was complementary to that of the receptor. The SLTs are AB toxins constituted by an enzymatic (A) component and a cell-binding (B) part. The A-subunit is situated on one face of the B-component, which is a pentamer of identical subunits. In the absence of the A-subunits, the B-subunits still form pentamers that are functionally equivalent to the overall toxin in their attachment to host cell. The *in vivo* cytotoxicity of SLTs has been correlated with their binding affinity toward the glycolipid

globotriaosylceramide (Gb_3). The crystal structure of SLT-I B-pentamer complexed with a Gb_3 analogue revealed three Gb_3-binding sites per B-subunit[113] (Figure 15).

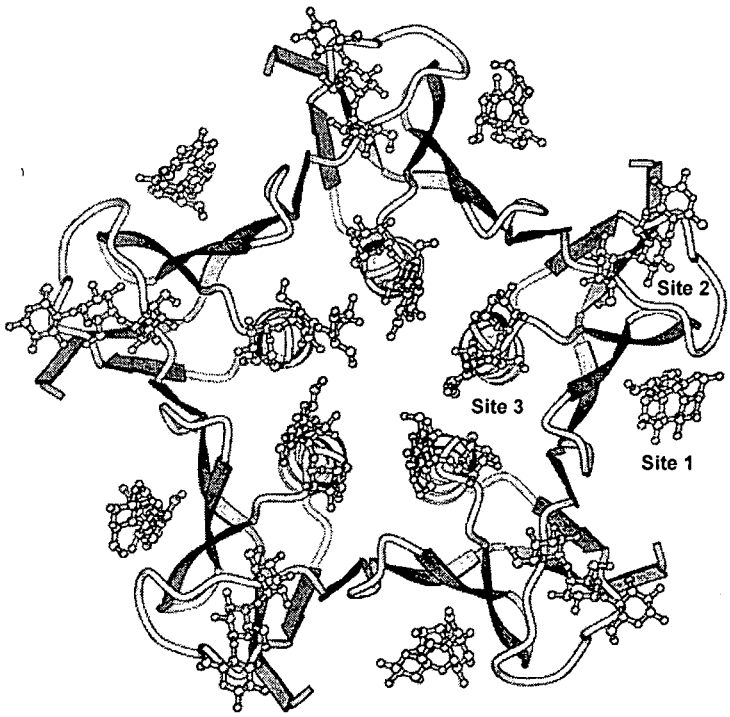


Figure 15 View along the 5-fold axis of the SLT-I B-pentamer bound to the Pk-MCO trisaccharides. The surface toward the viewer is the sugar binding surface.

As determined by means of an ELISA protocol, STARFISH exhibited more than 10^6-fold increase in inhibition over the Pk trisaccharide. STARFISH was designed to achieve high affinity through the simultaneous binding of all five peripheral bridged Pk dimers to the ten sites 1 and 2 of the B-pentamer (with each trisaccharide dimer bridging sites 1 and 2 of the monomer). Crystallographic studies of the formed complex revealed a different mode of binding (Figure 16). One STARFISH molecule, in fact, bound two B-subunit monomers from separate toxin molecules, with the saccharide residues occupying only the ten sites 2. Each trisaccharide interaction was identical to that seen for the corresponding univalent ligand. A higher affinity of binding of site 2 compared to 1 had been already highlighted by previous studies carried out on toxin mutants. Therefore, the observed mixture of aggregative and

chelate mechanisms through which STARFISH exerted its activity could be explained by the stronger interactions of the carbohydrate within site 2, possibly coupled with unfavourable entropic and enthalpic contributions of the bridging.

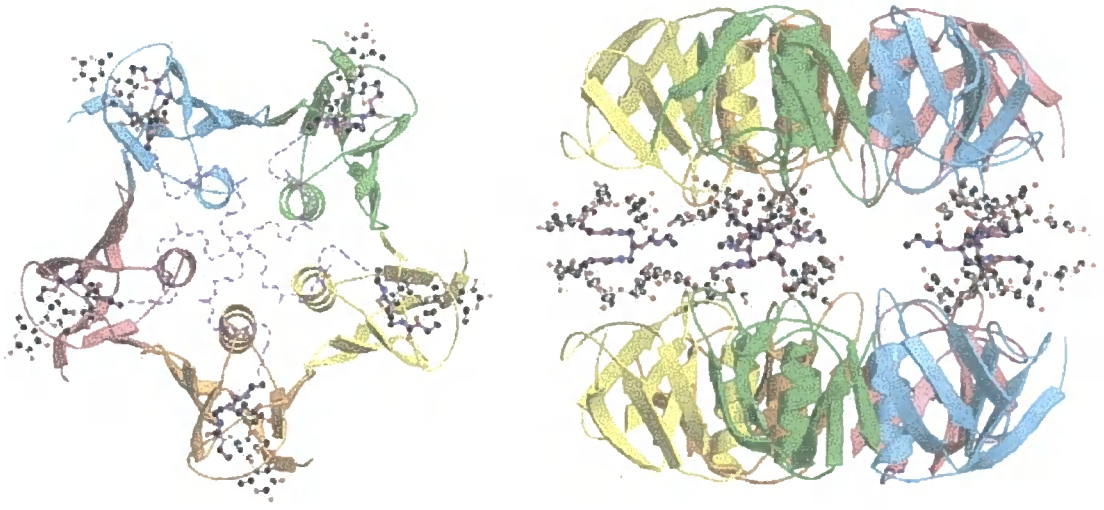


Figure 16 Diagrams of STARFISH ligand bound to the SLT.

A calorimetric study of the interaction of bivalent glycopeptides with SLT-I was carried out by Toone and co-workers[100]. The two ligands used differed in the nature of the peptidic linker. Both ligands showed affinity gain, but the mechanism through which this was achieved depended on the nature of the linker domain: while the compound characterised by a hydrophobic linker seemed to work through a chelate mechanism, the hydrophilic one bound by means of an intermolecular aggregative process. Again the binding enthalpy of the latter appeared to be significantly lower (in absolute value) than that of the monovalent compound, a trend that this group already recognised as the signature for an intermolecular binding. The affinity enhancement reached through the intramolecular mechanism was slightly higher than that of the hydrophilic compound. Peptidic spacers were chosen in the search for possible favourable additional interactions of the ligand with the protein surface. These interactions were given as an explanation for the higher affinity of the

ligand with the hydrophobic linker and for its intramolecular mechanism. For the hydrophilic linker, either there were no favourable contacts or there were in fact repulsive interactions of the linker with hydrophobic regions on the protein surface. Therefore, the affinity was lower and the binding proceeded via an intermolecular mechanism.

The greatest increases in activity on a valency-corrected basis are reported with polymeric ligands[98]. Nevertheless, the only polymeric ligands synthesised so far which seem to bind through an intramolecular mechanism are those prepared by Kanai *et al.*[114] (Figure 17). Several ligands with different backbone lengths were synthesised and tested as inhibitors for Con A-carbohydrate interaction. The maximum inhibitory activity seemed to coincide with an average length sufficient to span two lectin binding sites belonging to the same protein molecule, supporting the hypothesis of a chelate mechanism of binding.

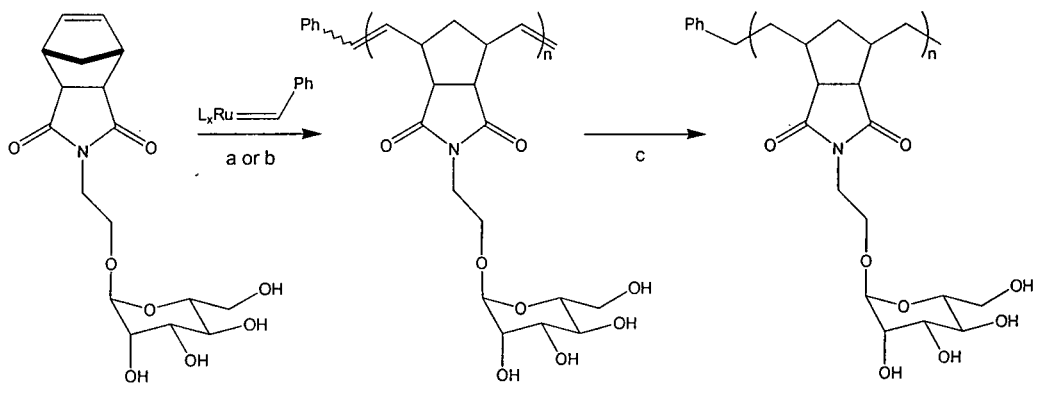


Figure 17 Scheme for the synthesis of neoglyopolymers by ROMP: (a), MeOH/H_2O/DCM (6:1:5) then add H_2O, room temperature; (b), H_2O/DCM (2:1), DTAB, room temperature; (c), $TsNHNH_2$, H_2O, 100 °C, 56% (n = 10), 91% (n = 25), 58% (n = 52), 100% (n = 143).

The enhancement in affinity determined for polymers too short to bridge two binding sites was attributed to a high local concentration of sugar moieties (statistical effect), which would perturb the rate of dissociation of the formed complex. The authors concluded that the observed dependence of the inhibition activity on the polymer

length was largely due to a combination of statistical and chelation effects. Moreover, despite the differences in backbone flexibility, the most potent ligands of each series, **2** and **3**, had approximately the same efficacy. This latter observation points out an interesting aspect of multivalent ligand-protein binding. It is believed that rigid linkers would favour the interaction, due to a reduction of the entropic penalties accompanying the binding[110,112,115]. However, the great enhancements achieved with flexible linkers and the negligible gain in activity observed with dendritic ligands characterised by rigid arms, have recently led Toone and co-workers to suggest that entropic concerns might be less severe than was previously thought and probably less severe than those related to unfavourable enthalpic contributions[98].

1.4 CARBOHYDRATE-BASED THERAPEUTICS

A myriad of physiological processes depend on the formation of receptor-ligand complexes at the cell surface. As already underlined, multivalent interactions are usually required. The ability to inhibit and/or activate, or exploit such interactions can bring enormous developments in medicinal chemistry. Consequently, extraordinary efforts are made in order to discover natural and non-natural products able to interfere with the cell surface binding events. In this respect, synthetic polyvalent ligands seem to represent the best option. Natural compounds are, in fact, often too scarce, too complex to identify the relevant molecular mechanisms of interaction and structurally heterogeneous[41]. When considering glycoconjugates, for example, unlike natural substances, synthetic materials, called neoglycoconjugates[116], possess the advantage of containing carbohydrate residues of known structure and assured purity. Between neoglycoconjugates, several glycopolymers represent probably the most promising materials, due to their lower

immunogenic and toxic effects compared, for example, to glycoproteins and to the possibility of bearing more sugar moieties and drug molecules than a protein of identical molecular mass[117]. In general, many features of multivalent synthetic ligands, such as the number and the orientation of the binding epitopes, the flexibility, size and shape of the scaffold from which the binding sites protrude, can be varied systematically, providing a powerful tool for the investigation of the single aspects of receptor-ligand complexation[118,119] and for the design of very effective compounds. Moreover, synthetic materials can either mimic natural products or alter the cellular function by mechanisms inaccessible to natural substances. Synthetic polyvalent ligands can be used to investigate the binding events, as diagnostic tools, to inhibit or activate receptor-ligand interactions and as drug carriers to deliver the pharmaceutical to a selected organ/tissue. Numerous kinds of specific receptor-ligand interactions are involved in the complex functioning of the mammalian body. Therefore, ligands containing different recognition moieties can be prepared.

The recognition of the vital role of lectin-carbohydrate interactions has provided the basis for the development of lectin- and carbohydrate-based therapeutics[120]. Lectins are nowadays employed as recognition determinants for the preparation of colloidal vehicles for, for example, the mucosal delivery of drugs and vaccines[121]. All the mucosal surfaces have the disadvantage that fluid secretion dilutes and may flush away applied delivery vectors. The presence of lectins improves the absorption of the delivery system: the lectin acts as a bioadhesive, interacting with specific carbohydrates expressed on the cell surface. The residence time of drugs and vaccines is so prolonged, reducing the dosage and the frequency of administration. Moreover, cell surface carbohydrate expression is cell type specific, allowing the targeting of the active agent to selected areas of the mucosal tract. One of the major

drawbacks for lectin-mediated mucosal delivery is represented by the possible cross-reactivity, where different areas presenting the same carbohydrate structure may compete with the selected cells for the lectin-binding. Furthermore, many lectins are toxic and/or immunogenic and the effects of repeated dosage are largely unknown. However, these latter problems could be overcome by using truncated forms of the proteins.

Carbohydrates are particularly suitable as recognition moieties. As already said, the combination of different monosaccharides can give potentially limitless oligosaccharide structures. This aspect, combined with their small size (*i.e.* a lower risk of immunogenity compared to large molecules) and their usually easy accessibility, makes them perfect candidates as targeting determinants. Neoglycoproteins have been shown to be potent drug carriers: muramyl dipeptide (MDP), an immunoactivator, bound to Man-BSA is able to activate macrophages to a tumoricidal state, reversing the development of lung metastases in mice[122]. However, concerns regarding immunogenicity and pI shifts of the carrier protein have moved the attention toward aglycons mainly derived by polymerisation procedures[123]. The possibility of obtaining biodegradable, non-toxic and non-immunogenic materials of known composition and with the required structural characteristics reproducibly and in high yields sets glycopolymers at the centre of the research for carbohydrate-based therapeutics.

Importantly, carbohydrates can be used not only as targeting moieties of drug delivery systems or inhibitors/activators, but also as scaffolds and drugs themselves. Sialic acid amino-acid (SAAs) oligomers have been synthesised to mimic natural polysaccharides, such as dextran sulphate and heparin, and tested as inhibitors of HIV replication. In this case, the use of synthetic ligands allows the overcoming of

problems of poor absorption, instability and anticoagulant activity associated with the natural products[124]. A macromolecular prodrug of cisplatin (CDDP, one of the most potent antitumour platinum complexes), has been recently synthesised by immobilising CDDP onto dextran (Dex) through a six-membered chelate-type coordination bond. The targetability of the product to hepatocytes was provided by introducing tetraantennary galactose units (Figure 18).

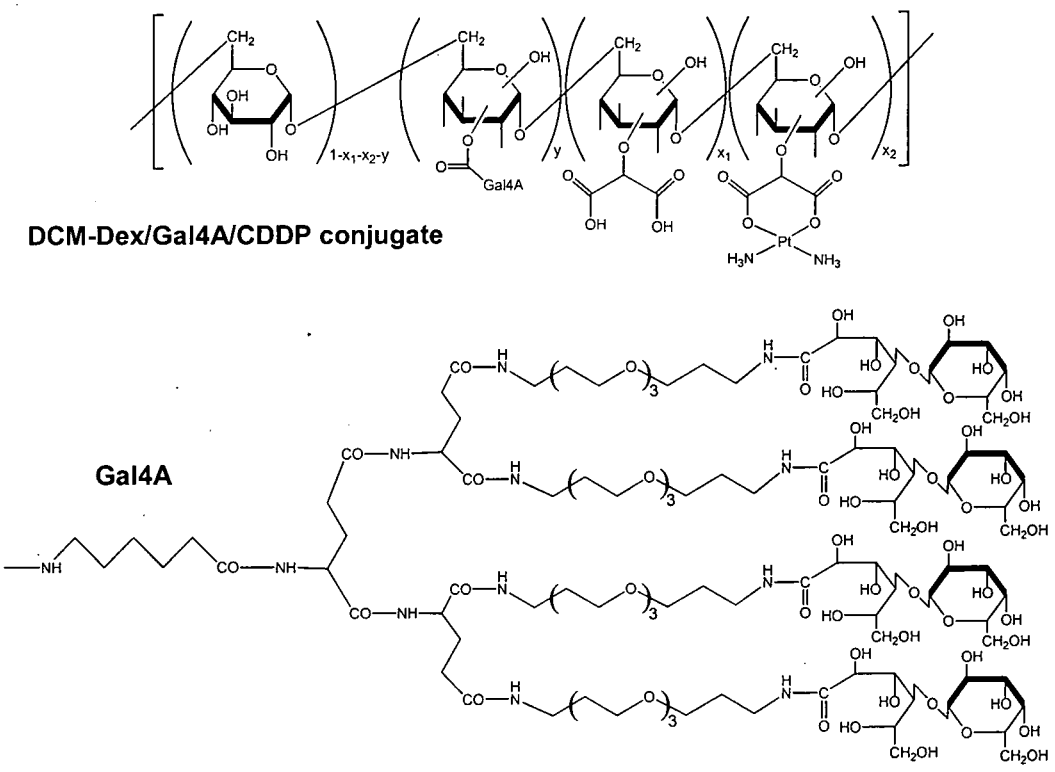


Figure 18 Structure of DCM-Dex/Gal4A/CDDP conjugate.

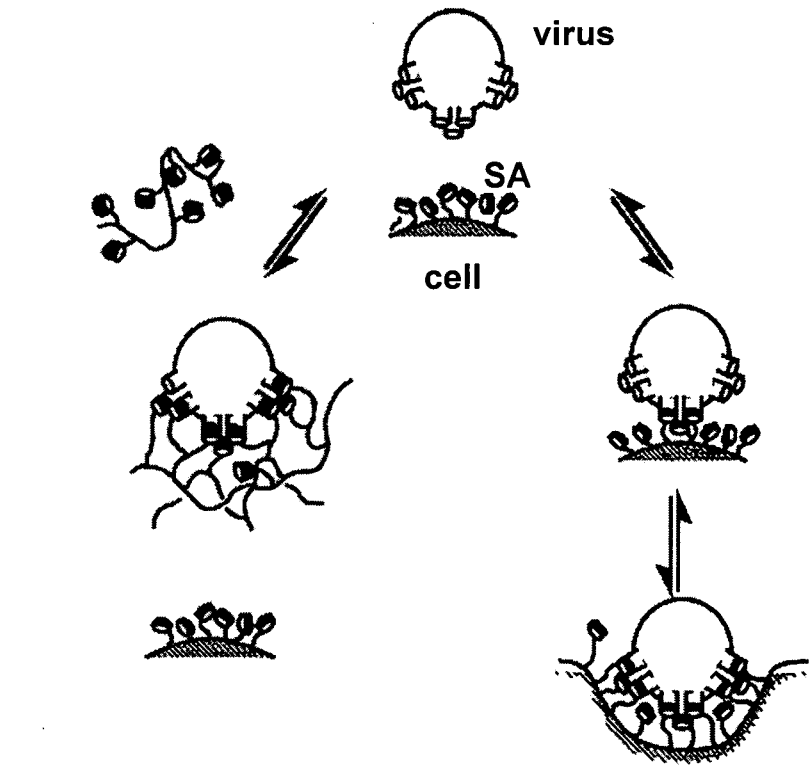
The binding constant of this conjugate, DCM-Dex/Gal4A/CDDP, to *Ricinus communis* agglutinin (RCA) was higher than that of dextran carrying monomeric galactose residues. Moreover, DCM-Dex/Gal4A/CDDP showed cell-specific cytotoxic activity against HepG2 *human hepatoma* cells *in vitro*. This activity was inhibited by the addition of galactose and Gal4A[125].

As for lectins[41,120], sugars can also be considered pharamaceutical agents in their own right. The rigidity of the pyranose ring allows the attachment of pharmacophoric groups with well-defined spatial orientation[124]. The same structural features can be exploited to introduce conformational constraints in pre-existing molecules[126]. Recent studies have shown that introduction of SAA in somatostatin (a 14-residue cyclic peptide hormone formed in the hypothalamus) leads to a pharmaceutical with strong antiproliferative activity against multi-drug resistant hepatoma carcinoma[127]. This is of special interest, since resistance to chemotherapy has become a serious problem in cancer therapy.

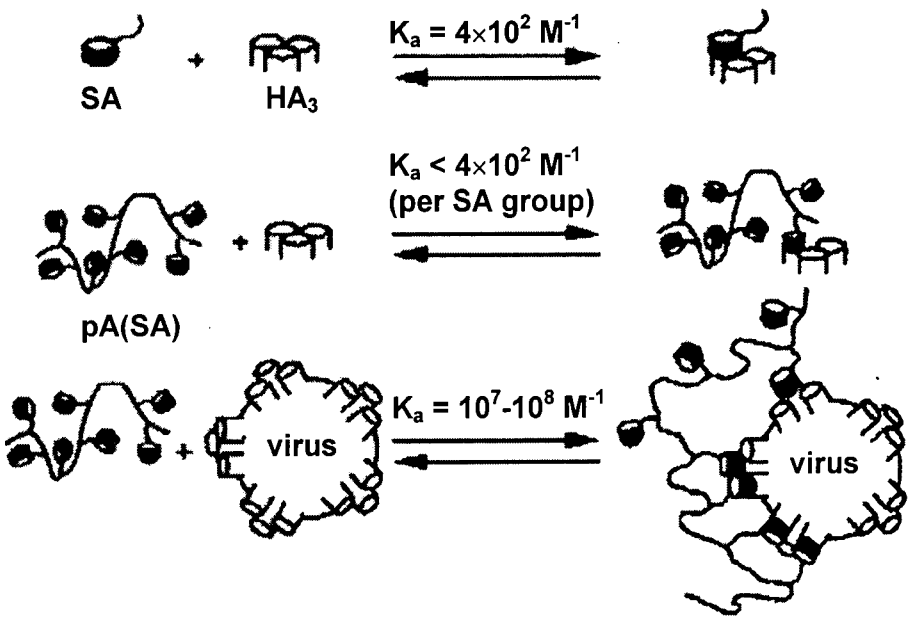
1.4.1 SYNTHETIC MULTIVALENT LIGANDS AS INHIBITORS AND EFFECTORS

Synthetic multivalent ligands can be used to explore/interfere with either cellular interactions in which the ligand prevents the binding (inhibitor) or in which it induces a response (effector).

Inhibitors can be employed to block viral and bacterial binding to the host cell and to inhibit the binding of leukocytes to endothelial cells. Some of the first applications of polyvalent inhibitors were focused on preventing the binding of influenza virus hemagglutinin to host cells[128] and this area of research is still active[129,130]. The virus binds the host cell through the interaction of its trimeric hemagglutinin with sialic acid residues on the cell surface (Figure 19a).



(a)



(b)

Figure 19 a), The attachment of the influenza virus to a target cell occurs through multiple simultaneous interactions between hemagglutin (HA) and sialic acid (SA). The attachment can be blocked by a polyvalent inhibitor; b), association constants of monomeric SA and HA, SA from polyvalent polyacrylamide and HA, and SA from polyvalent polyacrylamide and HA on the surface of the virus.

High concentrations of monovalent sialic acid derivatives are required in order to inhibit the binding (*ca.* 10^{-3} M)[131]. A variety of ligands bearing multiple residues of sialic acid has been synthesised, ranging from liposomes to glycopolymers. One of the most effective inhibitors, consisting of a linear polyacrylamide bearing sialic acid residues on the side chains, shows an association constant 10^5-10^6 times higher than the corresponding monomer[132] (Figure 19b). However, the concentration of polymer required to inhibit the cell-virus interaction is approximately 10^7-10^8 times lower than that of the monovalent ligand. The authors have suggested that approximately 10^6 of the total enhancement is due to an increased affinity of the polymer for the surface of the virus, whereas *ca.* 10^2 is due to steric stabilisation. The multivalent ligand forms a loose, water-swollen, gellike layer surrounding the viral surface, which sterically prevents the close approach of the interacting species[133].

Helicobacter pylori are the bacteria that cause gastric ulcers. They attach to gut cells by binding to extracellular sialylated glycoproteins. This adhesion has been effectively inhibited by albumin glycosylated with 3'-sialyllactosyl residues[134].

Again, studies have shown the critical importance of the ligand structure. Two examples of modular synthesis have been already described in Section 1.3.2.2[110,112]. The structures of the ligands are reported in Figure 20. In both cases, the extremely high inhibitory activity determined is due to the matching of the 3D-structure of the ligand with the respective toxin.

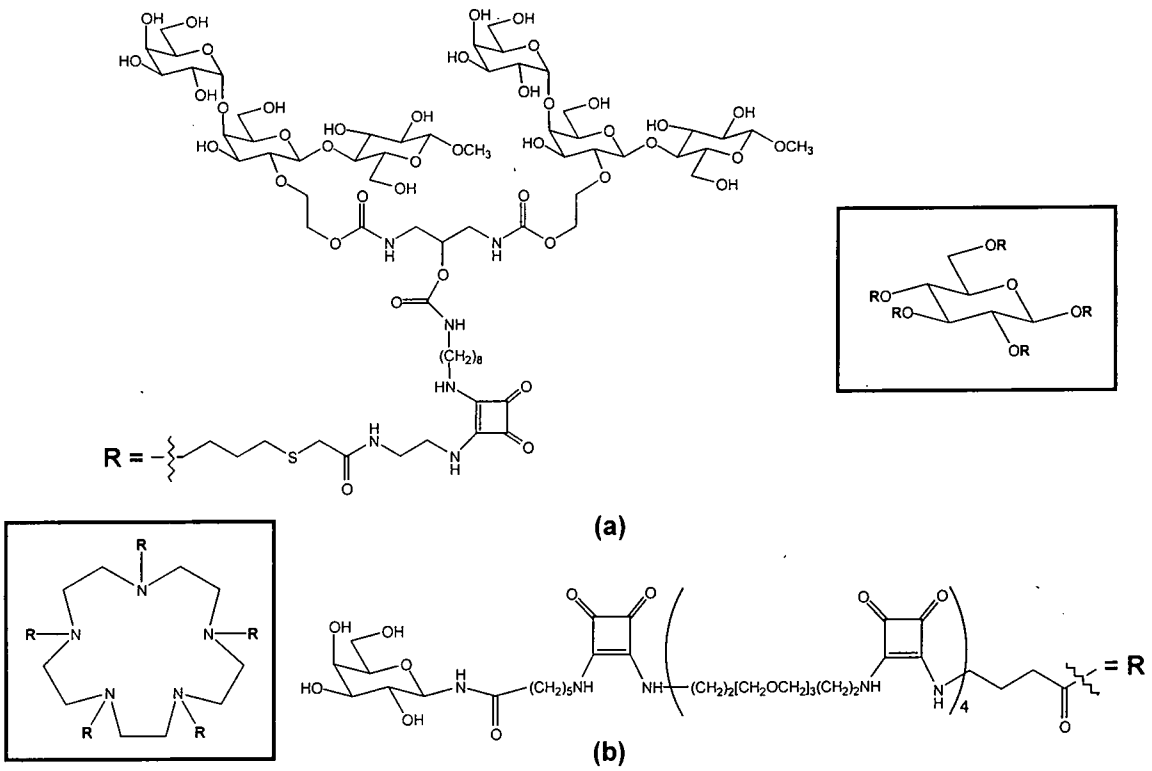


Figure 20 a), Structure of a decavalent ligand (STARFISH) that acts as a potent inhibitor of SLT-1[112]; b), structure of a pentavalent ligand that is an extremely effective inhibitor of LT-1[110].

Polyvalent ligands can also be used to manipulate signal transduction pathways. For instance, dimerisation of cell surface receptors induced by ligand binding can initiate a particular cellular response[135]. Divalent ligands have been used to investigate the role of receptor clustering in signal transduction[136]. The most common partially or fully synthetic multivalent activators are those that can cause an immune response. Examples are vaccines, which display multiple copies of an antigen bound to a synthetic scaffold or to a protein carrier. Great research is at present devoted to the formulation of carbohydrate-based anticancer vaccines[137,138]. In recent years, many tumour-associated carbohydrate antigens have been identified and some of them, including the Tn [GalNAcα(1$\rightarrow$*O*)Ser/Thr] and T [Galβ(1$\rightarrow$3)GalNAcα(1$\rightarrow$*O*)Ser/Thr] antigens, are highly expressed on carcinoma-

associated mucins, resulting from truncation of carbohydrate chains during malignant transformation[139]. St Hilaire and co-workers[140] have recently reported the synthesis of several glycopeptides containing the T-antigen potentially suitable as cancer vaccines. Glycopeptides with variations in both peptide amino acid sequence and distance between the T-antigen and the peptide scaffold were prepared in order to investigate the *in vivo* effects of these parameters. Glunz *et al.*[141] have recently reported the preparation of a hexasaccharide-keyhole limpet haemocyanin (KLH) protein conjugate. The hexasaccharide moiety, termed globo H, was originally isolated from a glycoceramide associated with breast cancer. The synthetic globo H-KLH conjugate was successfully used to induce high amounts of anti-globo H antibodies and caused cell lysis at levels approaching those of monoclonal antibodies raised against cancerous cells.

A polyvalent ligand may present simultaneously effectory and inhibitory properties. If the effector promotes receptor down-regulation, the same product may inhibit the receptor activity by means of the same mechanism[142]. Labeled glycopolymers have been recently synthesised by ring-opening metathesis polymerisation (ROMP) by Gordon and co-workers[143]. The materials were found to bind specifically L-selectin, a surface protein that mediates lymphocytes homing and leukocytes recruitment to inflammatory sites. The ligands act to down-regulate the amount of L-selectin at the cell surface[144] and at the same time to block L-selectin-mediated leukocyte rolling[145].

1.4.2 TARGETED DRUG DELIVERY

Targeted drug delivery arises from the field of cancer therapy, where chemotherapy involves the utilisation of cytotoxic drugs to kill cells. The majority of

pharmaceuticals currently in use have not been designed to exert their activity only at a certain disease site, resulting in an even distribution of the therapeutic agent throughout the body. The problem is that the activity of anticancer drugs against tumour and non-tumour cells is roughly the same, leading to serious damage of healthy tissues and, in some cases, to the death of the patient. The drug has, in fact, to cross many biological barriers before reaching the desired site of action, so affecting other organs. At the same time its concentration at the pathological area is significantly diminished. As a result, the dosage and the frequency of administration are largely augmented, increasing not only the life-threatening side effects, but also the complexity of the administration protocol and the expenses of the treatment. Drug targeting, that is the capability of delivering the therapeutic agent selectively and quantitatively to the organ affected by the pathology, represents a solution to all these problems. Ideally, targeted drug delivery should ensure a high drug concentration at the disease site, keeping that at non-target organs and tissues below acceptable minimal levels.

The concept of drug targeting was firstly suggested by Ehrlich roughly a century ago[117]. In this approach, drug targeting is achieved by means of a hypothetical "magic bullet" consisting of two components: one able to recognise and bind the target, the other providing the therapeutic action. Since the expression of some plasma membrane receptors and antigens is, to a first approximation, characteristic of certain cells or related to certain diseases, cell-specific targeting can be achieved using antibodies or receptor-specific ligands[146]. The first real attempt to target leukaemic cells via an antibody-linked methotrexate (MTX, one of the currently used anticancer drugs) dates back to 1958[147]. Since then enourmous progress has been made. The concept of drug targeting has been extended and, nowadays, the

"magic bullet" is generally thought as the combination of three elements: the drug, the targeting moiety and a macromolecular carrier used to increase the residence time of the therapeutic agent in the body and to multiply the number of drug molecules per targeting unit[148].

1.4.2.1 TYPES OF DRUG TARGETING

Most of the drug delivery systems prepared so far lack the targeting moiety. In these cases the actual targeting is achieved by the exploitation of differences between healthy and tumour tissues[149]. Thus, especially for the treatment of solid tumours, a certain distribution control can be attained by taking advantage of the enhanced vascular permeability of tumour tissues[150]. The leaky vasculature of neoplastic regions can, in fact, allow the extravasation of even relatively large particles, with diameters ranging from 10 to 500 nm. As a result, high molecular weight substances carrying the therapeutic agent preferentially accumulate at these sites. This is a type of passive targeting, known as "Enhanced Permeability and Retention Effect, EPR"[151-154].

It is also known that inflamed and neoplastic areas usually show some acidosis and hyperthermia. This makes possible the use of stimuli-responsive drug carriers that can disintegrate at lower pH values or at higher temperatures, releasing the drug only at the affected sites[155,156]. pH-sensitive liposomes, such as 3β-(*N*-(*N'*,*N'*-dimethylaminoethane)carbamoyl)cholesterol:dioleoylphosphatidylethanolamine (DC-Chol:DOPE), have been shown to give good transgene expression in both epithelial cells *in vitro*[157] and directly injected human tumours in mice[158]. Several temperature-responsive block copolymer micelles have been prepared using poly(*N*-isopropylacrylamide), PNIPAAm. Above the lower critical solution temperature

(LCST), PNIPAAm becomes insoluble with subsequent destabilisation of the micelles and burst-like release of the entrapped drug[159,160].

Despite the good results obtained with these drug targeting methodologies, often the affected area does not differ much from normal tissues in terms of vascular permeability, temperature and acidity and the targeting strategies described above may become ineffective. The most efficient way to impart specificity toward the target organ to a drug-macromolecule conjugate is to bind it to another molecule able to recognise specifically and bind the target site. A wide range of substances can be used as targeting moieties: antibodies and their fragments, lectins, lipoproteins, hormones, saccharides, peptides and folate[149].

Monoclonal antibodies are the most employed targeting molecules. Nevertheless, besides the problems of immunogenicity associated with their use, other aspects are limiting their suitability as recognition determinants. First of all, problems of steric hindrance can complicate the attachment of the antibody to the water-exposed termini of the hydrophilic block of block copolymer micelles used as drug carriers[161], or the coupling of the antibody with the pharmaceutical agent can result in the loss of its specificity[162]. Moreover, even when these problems are not encountered, antibody-carrying species may still be captured by circulating antigens before reaching the target[162] and, most important, most of the "tumour-associated" antigens are only quantitatively different in tumour and normal tissues, leading to serious problems of cross-reactivity[163].

The reasons that make carbohydrates good candidates as targeting moieties have been already discussed throughout the previous sections.

1.4.2.2 RECEPTOR-MEDIATED ENDOCYTOSIS, RME

The ligands present on the surface of the carrier molecules can mediate cellular uptake by interacting with specific receptors expressed on the surface of the target cells. This initial interaction may induce the internalisation of the drug-carrier conjugate by a process called receptor-mediated endocytosis, RME[164,165]. The concept of RME was formulated in 1974 by Ashwell and Morell[48], and Goldstein and Brown[166] to explain the hepatic uptake of asialoglycoproteins (ASGPs) and low-density lipoproteins (LDLs), respectively. A simplified schematisation of RME is reported in Figure 21.

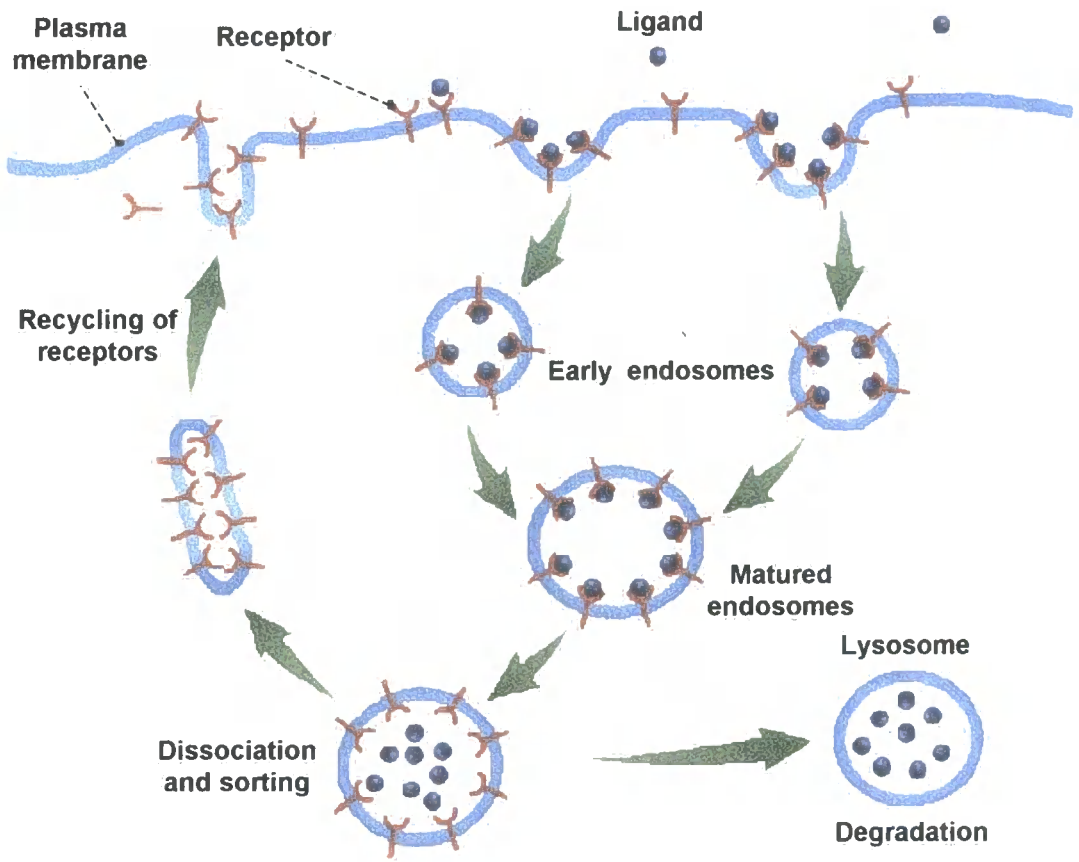


Figure 21 Receptor-mediated endocytosis, RME.

The process is initiated by the interaction of a cell surface receptor with its specific ligand present, for example, on the surface of the drug carrier. The receptor (-ligand complex) moves to and clusters within special regions of the plasma membrane termed coated pits. These invaginations of the plasma membrane, approximately 100 nm in diameter, have a cytoplasmic coat composed of clathrin (see Figure 8). Coated pits are then excised from the membrane to form coated vesicles. These rapidly lose their clathrin shells and fuse with endosomes. In turn, endosomes fuse with one another to form mature endosomes having diameters ranging from 200 to 600 nm. The slightly acidic pH of the endosomes (pH 6.3-6.8) induces dissociation of the receptor-ligand complex. Whilst the receptor is recycled and reinserted into the plasma membrane, the ligand-carrier conjugate is generally directed to the more acidic lysosomal compartment (pH 4.5-5.0), rich in enzymes, where the carrier is degraded and the drug released. Low-molecular weight pharmaceuticals can then cross the lysosome membrane and exert their action.

Clathrin-mediated uptake is believed to be the principal mode of cellular endocytic internalisation[163]. Moreover, the presence of lectins in endosomes and in the nucleus[167] suggests that lectin-carbohydrate interactions could be exploited not only for cellular but also for sub-cellular targeting. However, it should be noted that most of the receptors are not entirely specific, since they are expressed at different levels on different tissues. Nevertheless, receptor-mediated endocytosis of carbohydrate-carrying drug delivery systems in the liver represents one of the few examples in which quite selective drug targeting has been achieved *in vivo*[165].

1.4.2.3 MACROMOLECULAR CARRIER

Numerous types of drug carriers have been employed, including proteins and glycoproteins, soluble polymers, micro- and nanoparticles made of insoluble or biodegradable natural and synthetic polymers, microcapsules, cells, liposomes and micelles[122,151,161,168-178]. Each of these carrier types shows its own advantages and disadvantages and the choice of a certain structure has to be made only after a thorough consideration of all the aspects of the case to which it has to be applied.

Since 1972, when Gregoriadis proposed the use of liposomes as carriers of enzymes in the treatment of lysosomal storage disease, the application of liposomes has been extended to a variety of drugs[179,180]. Liposomes can be used to entrap hydrophobic drugs, such as most of the anticancer therapeutic agents currently in use, increasing their bioavailability, their residence time in the body and protecting them from blood components whose action could result in their degradation and inactivation[181]. However, following intravenous administration, liposomes tend to be rapidly removed from the circulation by the phagocytic cells of the reticuloendothelial system (RES)[182]. Phagocytosis by elements of the RES in the liver is mediated by adsorption of opsonins on the particle surface. Hydrophilically stabilised nanoparticles can avoid phagocytosis and achieve prolonged circulation times[183] (Figure 22). Thus, modification of liposome surfaces with poly(ethylene glycol) is often employed[184]. Glycosylation of liposomes can be attained through coating with glycoproteins or by incorporation of synthetic glycolipids[185]. However, the introduction of glycoproteins is rather difficult and problems of reproducibility and immunogenicity are often encountered. On the other hand, low-molecular weight glycolipids can be easily removed from the liposome via interaction with lipoproteins. Furthermore, the relatively large size of liposomes make them unusable when certain

areas have to be reached: for hepatocytes the dimension of the drug carrier should not exceed 150 nm in diameter[185].

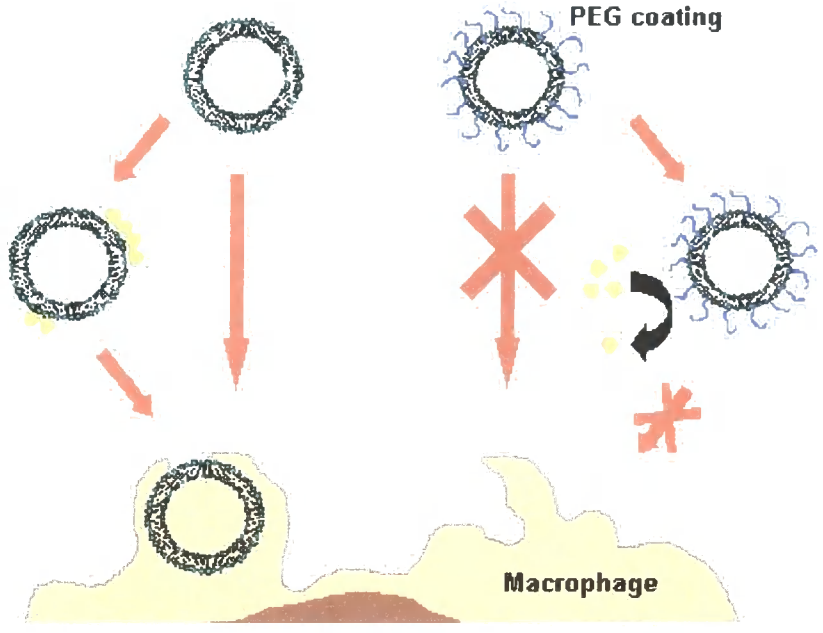


Figure 22 Modification of the liposome surface by PEG prevents its capture by the RES (orange spheres represent opsonins).

The use of partially or fully synthetic glycoproteins and glycopolymers allows the overcoming of some of these problems, especially those related to size restrictions and, obviously, those concerning the targeting moiety. Glycoproteins have been successfully used as drug carriers[62]. Complexes of albumin with oligosaccharides have been employed in experimental chemotherapy as vectors of anticancer agents. However, their clinical use is limited by an immune response that they might cause[168].

One class of drug delivery vehicles that has and is receiving great attention is represented by micelles formed by self-assembly of amphiphilic block copolymers[186.187]. It has been shown that amphiphilic AB-type block copolymers with the length of the hydrophilic block exceeding to some extent that of the

hydrophobic one can form spherical micelles in aqueous media[188]. The aggregates are composed of an inner hydrophobic core stabilised by a corona of hydrophilic polymer chains (Figure 23).

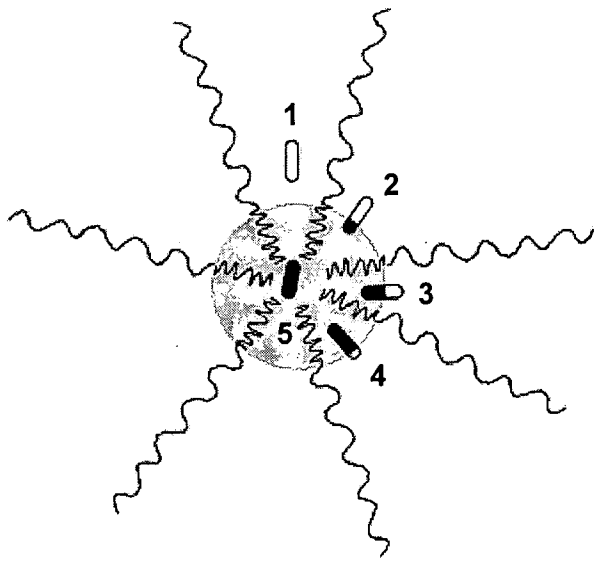


Figure 23 Structure of a micellar drug carrier. The pattern of drug association with the micelle depends on the drug hydrophobicity (black colour of the drug molecule shows hydrophobic area, white shows the hydrophilic one). Completely water-soluble hydrophilic drug can only be adsorbed within the micelle corona (case 1); completely insoluble hydrophobic molecules can only be incorporated into the micelle core (case 5); drug molecules with intermediate hydrophobic/hydrophilic ratio will have intermediate positions within the carrier particle (cases 2-4).

This core-shell architecture is ideally suitable for targeted drug delivery. First of all, a significant amount of a variety of drugs, including genes and proteins, can be loaded into the aggregate and delivered to the target site. For hydrophobic drugs, the entrapment into micelles' core increases their bioavailability, reduces their toxicity and other associated side effects, protects them from the surrounding biological milieu and may enhance their permeability across physiological barriers[189].

The distribution of the drug-loaded micelles in the body is mainly determined by their size and their surface properties. Both these parameters can be controlled, for instance varying the length of the hydrophobic[190] and hydrophilic[191] blocks and

differently functionalising the hydrophilic segment. The hydrophobic core is generally constituted by biodegradable polymers, such as poly(lactic acid) (PLA)[192], poly(lactic-*co*-glycolic acid) (PLGA)[193] or poly(ε-caprolactone)[194]. The hydrophilic shell is usually composed of nonbiodegradable biocompatible polymers. Poly(ethylene glycol) (PEG) of molecular weight from 1 to 15 kDa remains the most widespread choice, due to its high solubility, high hydration and efficiency to act as steric protector in biological media[161]. The presence of PEG, in fact, increases the aggregates longevity, suppressing their non-specific uptake by the RES (see Figure 22).

Alternatively, the starting copolymers can be prepared from two hydrophilic blocks, one of which is then modified by covalent attachment of or non-covalent interaction with a hydrophobic pharmaceutical (cisplatin, adriamycin, doxorubicin, hydrophobic diagnostic units, charged drugs, DNA, etc.), yielding an amphiphilic micelle-forming polymer[195]. Kataoka and co-workers have used this procedure to conjugate several anticancer drugs (cisplatin[171], adriamycin[153,196] and doxorubicin[170,197]) to the poly(aspartic acid) (PAA) block of PEG-PAA micelles. The structure of doxorubicin-conjugated PEG-PAA block copolymer is reported in Figure 24. It is noteworthy that further loading of the drug can be attained by physical entrapment of the therapeutic agent into the inner core. The physical incorporation of doxorubicin results in an increase of the micelles' stability and in a very high *in vivo* antitumour activity. PEG-PAA(DOX) micelles, with both chemically bound and physically entrapped doxorubicin, were shown to achieve prolonged circulation time in the blood and to accumulate significantly at the solid tumour by exploitation of the EPR effect. The complete tumour regression observed was attributed mainly to a sustained release of physically entrapped doxorubicin.

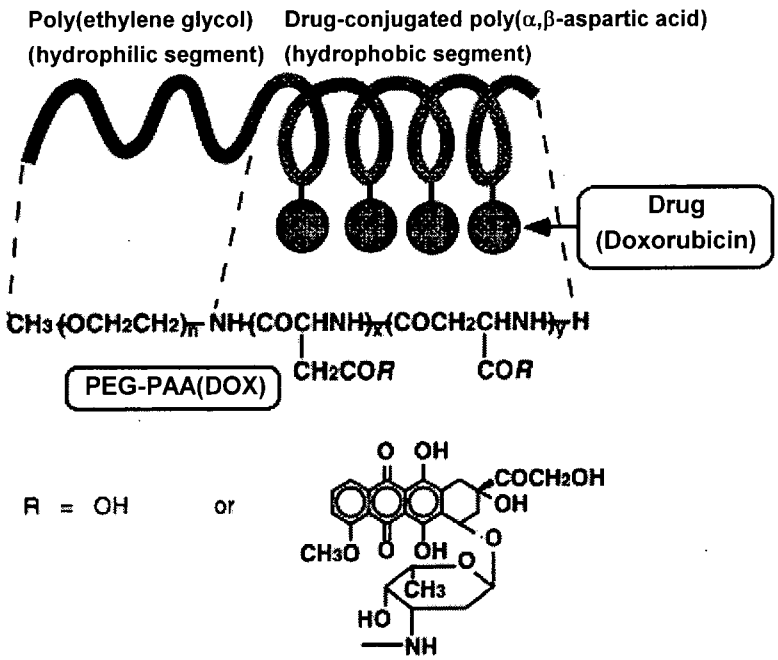


Figure 24 Structure of doxorubicin-conjugated poly(ethylene glycol)-poly(α,β-aspartic acid) block copolymer.

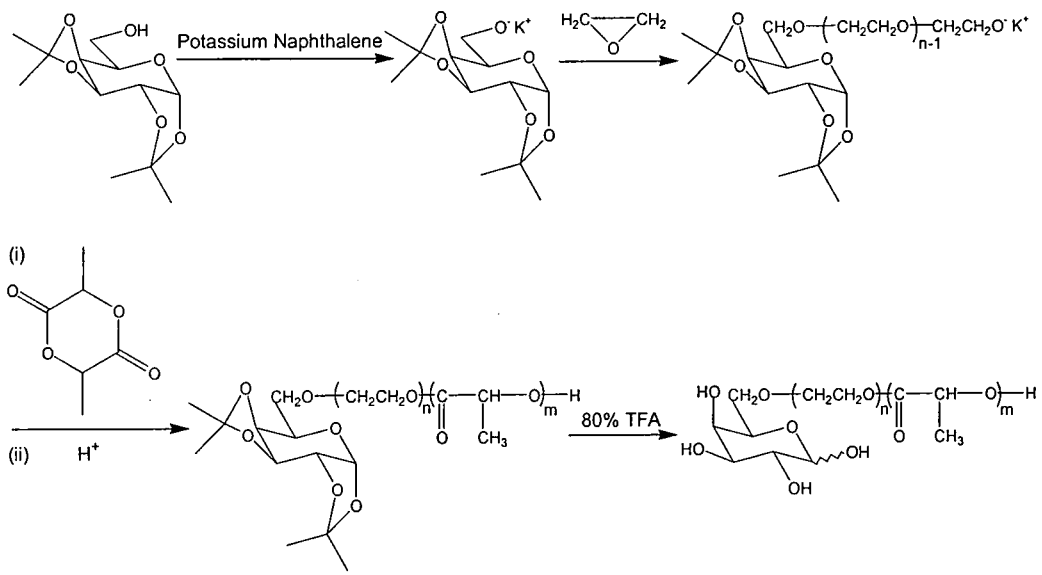
A simpler system has been developed by the same authors, where the micelle-forming copolymer is poly(ethylene glycol)-poly(β-benzyl-l-aspartate) (PEG-PBLA). Doxorubicin was successfully incorporated within the core, due the structure matching of the drug and the hydrophobic core (π-π interactions). Micelle-entrapped doxorubicin was shown to have considerably higher antitumour activity than the free drug[198].

All the micelles described above lack a recognition moiety. Their targeting to the pathological area is achieved only through the spontaneous accumulation in regions presenting a leaky vasculature.

Amphiphilic block copolymers of vinyl ethers having pendant *N*-acetyl-D-glucosamine residues were synthesised by Yamada and co-workers[199]. The interaction of the polymer with wheat germ agglutinin (WGA) was investigated and the association constants evaluated by measuring the changes in fluorescence of WGA

upon binding. It was demonstrated that the GlcNAc-carrying polymer shows a significant increase in binding affinity toward the lectin compared to GlcNAc and its oligomers.

Kataoka *et al.*[200] have recently reported the synthesis of poly(ethylene glycol)-poly(d,l-lactide) block copolymers containing galactose residues at the PEG chain end. The copolymers were obtained by successive anionic ring-opening polymerisation of ethylene oxide and d,l-lactide using a metalated protected sugar as initiator (Scheme 1).



Scheme 1 Synthetic route of galactose-PEG-PLA block copolymers by anionic ring-opening polymerisation.

The presence of galactose residues at the surface was confirmed by passing the micelles loaded with pyrene through a column packed with RCA lectin immobilised beads. The galactose-containing micelles were retained in the column, indicating a strong lectin-carbohydrate interaction.

The possibility of physically entrapping hydrophobic drugs into the inner core represents a great advantage of micellar drug delivery systems. In fact, the covalent

binding of pharmaceuticals can result in the loss of their activity once they are released from the carrier. Moreover, the macromolecule-drug bond needs to be accessible so that it can be cleaved. Therefore, linkers are usually required. The best option is represented by linkers which can be cleaved only intracellulary by specific enzymes[163]. Spacers of this type contain a variety of amino acids. Several peptide linkers have been coupled to the carboxyl groups of methotrexate (MTX). The rate of release of the drug has been shown to depend on the length of the spacer[163].

Polymeric micelles are generally more stable than surfactant ones, showing a significantly lower critical micelle concentration (CMC) and a slower rate of dissociation[201]. Therefore, polymeric micellar aggregates may retain their structure upon dilution after injection, allowing retention of the loaded drug and, eventually, its accumulation at the target site. In principle, a micellar drug delivery system should be stable enough to reach the pathological area and once there allow the release of the active agent. The stability of micelles can be increased by cross-linking of either the core or the corona[202,203]. Kataoka *et al.*[202] have recently reported the preparation of reversibly cross-linkable micelles constituted of poly(ethylene glycol)-poly(l-lysine) (PEG-PLL) by introduction of thiol groups on a small fraction of the lysine repeating units. The soluble copolymer is made amphiphilic by addition of DNA. Oxidation of the thiol groups results in the formation of disulfide bonds and cross-linking of the core. Disulfide bonds can be easily broken by an appropriate reducing agent. The most abundant reducing agent in cells is glutathione (GSH). The intracellular GSH concentration is 300-fold higher than that in the blood. The authors demonstrated that the micelles, called polyion complex micelles (PICs), are stable at low concentrations of GSH, whereas they dissociate increasing the concentration of the reducing agent at the millimolar level, as present in cells (Figure 25). These experiments clearly

demonstrated the potential of block copolymer micelles as carriers for DNA in gene therapy.

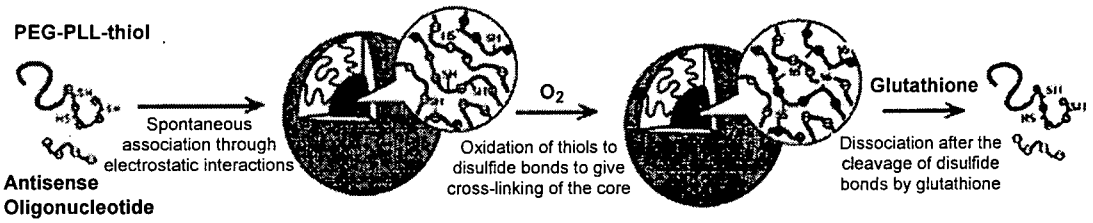


Figure 25 Schematic representation of the formation of reversibly cross-linked micelles from thiol-modified PEG-PLL block copolymers.

Cross-linking of the corona could lead not only to highly stable micelles, but subsequent removal of a degradable inner core could allow the preparation of nanosized hollow particles[204,205]. Such nanocontainers might be of particular interest for drug delivery, due to their higher loading capacity compared to the original micelles and due to the fact that they might provide an interior suitable for the transport of hydrophilic molecules. Although shell cross-linked block copolymer micelles and their corresponding nanocages have been developed only recently, initial studies have suggested that they posses great potential for the preparation of advanced drug delivery systems[206,207].

1.5 CONCLUDING REMARKS

Recent progress in glycobiology has revealed that cell surface oligosaccharides play an essential role in recognition events. The ubiquity of lectin-carbohydrate interactions throughout biological processes opens enormous potential for their exploitation in medicine. Lectin-carbohydrate associations are usually weak (K of the order of mM for monosaccharides) when compared to others in Nature. However, when more saccharides of the right type are clustered together with the

right geometry, the interaction becomes strong and highly specific. This increase is more than what would be expected on the basis of the increased sugar local concentration. The reasons of this phenomenon, termed "cluster glycoside effect", have not been rigorously determined yet, but its implications are extraordinary.

Lectins are good candidates as model systems for elucidating protein-carbohydrate interactions. On the other hand, neoglycoconjugates seem to represent the most suitable counterpart. Unlike natural products, they contain sugar residues of known structure, they are homogeneous and they can be prepared reproducibly and in large quantities.

It has been repeatedly pointed out that a thorough understanding of the lectin-carbohydrate interaction is absolutely required if this is wanted to be exploited for biomedical applications[62,66,162]. A deep knowledge of both structural and energetic aspects of the process is thus necessary. Indeed, the failure of a number of carbohydrate-based drugs may be attributed to a poor understanding of their supposed mechanism of action rather than to any inherent deficiency of carbohydrate therapeutics[62]. It has, in fact, been shown that when good understanding was achieved, the results have been impressive[110,112].

1.6 BIBLIOGRAPHICAL REFERENCES

[1] W. C. Boyd, *Science* **1954**, *119*, 419-426.

[2] G. W. G. Bird, *Vox Sang.* **1959**, *4*, 307-313.

[3] W. C. Boyd, *Vox Sang.* **1963**, *8*, 1-32.

[4] N. Sharon, H. Lis, *Science* **1972**, *177*, 949-959.

[5] J. Kocourek, in *The Lectins: Properties, Functions and Applications in Biology and Medicine* (Eds.: I. E. Liener, N. Sharon, I. J. Goldstein), Academic Press, Orlando, **1986**, pp. 1-32.

[6] N. Sharon, H. Lis, *Science* **1989**, *246*, 227-234.

[7] I. E. Liener, N. Sharon, I. J. Goldstein, *The Lectins: Properties, Functions and Applications in Biology and Medicine*, Academic Press, Orlando, **1986**.

[8] N. Sharon, H. Lis, *Lectins*, Chapman and Hall, London, **1989**.

[9] H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637-674.

[10] I. J. Goldstein, R. D. Poretz, in *The Lectins: Properties, Functions and Applications in Biology and Medicine* (Eds.: I. E. Liener, N. Sharon, I. J. Goldstein), Academic Press, Orlando, **1986**, pp. 33-247.

[11] R. D. Poretz, I. J. Goldstein, *Biochem. Pharmacol.* **1971**, *20*, 2727-2739.

[12] A. Pusztai, *Plant Lectins*, Cambridge University Press, Cambridge, **1991**.

[13] Y. D. Lobsanov, M. A. Gitt, H. Leffler, S. H. Barondes, J. M. Rini, *J. Biol. Chem.* **1993**, *268*, 27034-27038.

[14] N. K. Sauter, J. E. Hanson, G. D. Glick, J. H. Brown, R. L. Crowther, S.-J. Park, J. J. Skehel, D. C. Wiley, *Biochemistry* **1992**, *31*, 9609-9621.

[15] K. Drickamer, *J. Biol. Chem.* **1988**, *263*, 9557-9560.

[16] I. Ofek, R. J. Doyle, *Bacterial Adhesion to Cells and Tissues*, Chapman and Hall, London, **1994**.

[17] Y. Konami, K. Yamamoto, T. Osawa, T. Irimura, *Glycoconjugate J.* **1995**, *12*, 128-134.

[18] J. B. Sumner, S. F. Howell, *J. Bacteriol.* **1936**, *32*, 227-237.

[19] M. Inbar, L. Sachs, *Nature* **1969**, *223*, 710-712.

[20] N. Srinivasan, S. D. Rufino, M. B. Pepys, S. Wood, T. L. Blundell, *Chemtracts Biochem. Mol. Biol.* **1996**, *6*, 149-164.

[21] V. R. Srinivas, G. B. Reddy, N. Ahmad, C. P. Swaminathan, N. Mitra, A. Surolia, *Biochim. Biophys. Acta* **2001**, *1527*, 102-111.

[22] W. I. Weis, K. Drickamer, *Annu. Rev. Biochem.* **1996**, *65*, 441-473.

[23] B. G. Davis, *J. Chem. Soc., Perkin Trans. 1* **1999**, 3125-3237.

[24] J. M. Rini, *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 551-577.

[25] E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*, 719-728.

[26] R. Adar, N. Sharon, *Eur. J. Biochem.* **1996**, *239*, 668-674.

[27] A. Imberty, F. Casset, C. V. Gegg, M. E. Etzler, S. Pèrez, *Glycoconjugate J.* **1994**, *11*, 400-413.

[28] N. Sharon, in *Lectin Blocking. New Strategies for the Prevention and Therapy of Tumor Metastasis and Infectious Diseases* (Eds.: J. Beuth, G. Pulverer), Verlag, Stuttgart, **1994**, p. 3.

[29] N. M. Young, R. P. Oomen, *J. Mol. Biol.* **1992**, *228*, 924-934.

[30] V. Sharma, A. Surolia, *J. Mol. Biol.* **1997**, *267*, 433-445.

[31] M. E. Etzler, in *The Lectins: Properties, Functions and Applications in Biology and Medicine* (Eds.: I. E. Liener, N. Sharon, I. J. Goldstein), Academic Press, Orlando, **1986**, pp. 371-435.

[32] W. J. Peumans, E. J. M. Van Damme, in *Lectins. Biomedical Perspectives* (Eds.: A. Pusztai, S. Bardocz), Taylor & Francis, London, **1995**, pp. 1-21.

[33] J. W. Kijne, *Chemtracts Biochem. Mol. Biol.* **1996**, *6*, 180-193.

[34] P. Sears, C.-H. Wong, *Cell. Mol. Life Sci.* **1998**, *54*, 223-252.

[35] A. Varki, *Glycobiology* **1993**, *3*, 97-130.

[36] R. A. Laine, *Glycobiology* **1994**, *4*, 759-767.

[37] C. A. Bush, *Curr. Opin. Struct. Biol.* **1992**, *2*, 655-660.

[38] Y. Ni, I. Tizard, *Vet. Immunol. Immunop.* **1996**, *55*, 205-223.

[39] M. J. Kiefel, M. von Itzstein, *Chem. Rev.* **2002**, *102*, 471-490.

[40] D. Spillmann, *Glycoconjugate J.* **1994**, *11*, 169-171.

[41] N. Yamazaki, S. Kojima, N. V. Bovin, S. André, S. Gabius, H.-J. Gabius, *Adv. Drug Deliv. Rev.* **2000**, *43*, 225-244.

[42] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683-720.

[43] J. D. Rawn, *Biochimica*, Mc Graw-Hill, Milano, **1989**.

[44] F. M. Burnet, *Physiol. Rev.* **1951**, *31*, 131-150.

[45] R. B. Parekh, R. A. Dwek, J. R. Thomas, T. W. Rademacher, G. Opdenakker, A. J. Wittwer, S. C. Howard, R. Nelson, N. R. Siegel, M. G. Jennings, N. K. Harakas, J. Feder, *Biochemistry* **1989**, *28*, 7644-7662.

[46] G. E. Ritchie, B. E. Moffatt, R. B. Sim, B. P. Morgan, R. A. Dwek, P. M. Rudd, *Chem. Rev.* **2002**, *102*, 305-319.

[47] J. Roth, *Chem. Rev.* **2002**, *102*, 285-303.

[48] G. Ashwell, J. Harford, *Annu. Rev. Biochem.* **1982**, *51*, 531-554.

[49] Y. C. Lee, C. P. Stowell, M. J. Krantz, *Biochemistry* **1976**, *15*, 3956-3963.

[50] E. M. Anders, C. A. Hartley, D. C. Jackson, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 4485-4489.

[51] A. J. Tenner, S. L. Robinson, R. A. B. Ezekowitz, *Immunity* **1995**, *3*, 485-493.

[52] W. I. Weis, *Structure* **1994**, *2*, 147-150.

[53] M. Ohta, M. Okada, I. Yamashina, T. Kawasaki, *J. Biol. Chem.* **1990**, *265*, 1980-1984.

[54] I. Ofek, J. Goldhar, Y. Keisari, N. Sharon, *Annu. Rev. Microbiol.* **1995**, *49*, 239-276.

[55] T. Feizi, M. Larkin, *Glycobiology* **1990**, *1*, 17-33.

[56] M. Sumiya, M. Super, P. Tabona, R. J. Levinsky, T. Arai, M. W. Turner, J. A. Summerfield, *Lancet* **1991**, *337*, 1569-1570.

[57] K. Yamamoto, C. Ishida, Y. Shinohara, Y. Hasegawa, Y. Konami, T. Osawa, T. Irimura, *Biochemistry* **1994**, *33*, 8159-8166.

[58] L. A. Lasky, *Annu. Rev. Biochem.* **1995**, *64*, 113-139.

[59] H. Lis, N. Sharon, *Curr. Opin. Struct. Biol.* **1991**, *1*, 741-749.

[60] S. Teneberg, P. Willemsen, F. K. De Graaf, K.-A. Karlsson, *FEBS Lett.* **1990**, *263*, 10-14.

[61] Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, *28*, 321-327.

[62] B. G. Davis, *Chem. Rev.* **2002**, *102*, 579-601.

[63] M. N. Janakiraman, C. L. White, W. G. Laver, G. M. Air, M. Luo, *Biochemistry* **1994**, *33*, 8172-8179.

[64] P. Adler, S. J. Wood, Y. C. Lee, R. T. Lee, J. W. A. Petri, R. L. Schnaar, *J. Biol. Chem.* **1995**, *270*, 5164-5171.

[65] Z. K. Indik, J. G. Park, S. Hunter, A. D. Schreiber, *Blood* **1995**, *86*, 4389-4399.

[66] M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2754-2794.

[67] D. M. Sigal, J. D. Taurog, H. Metzger, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 2993-2995.

[68] M. C. Chervenak, E. J. Toone, *J. Am. Chem. Soc.* **1994**, *116*, 10533-10539.

[69] B. A. Williams, M. C. Chervenak, E. J. Toone, *J. Biol. Chem.* **1992**, *267*, 22907-22911.

[70] M. Kaul, D. S. Pilch, *Biochemistry* **2002**, *41*, 7695-7706.

[71] K. A. Watson, E. P. Mitchell, L. N. Johnson, J. C. Son, C. J. F. Bichard, M. G. Orchard, G. W. J. Fleet, N. G. Oikonomakos, D. D. Leonidas, M. Kontou, A. Papageorgioui, *Biochemistry* **1994**, *33*, 5745-5758.

[72] F. P. Schwarz, K. Puri, A. Surolia, *J. Biol. Chem.* **1991**, *266*, 24344-24350.

[73] F. P. Schwarz, K. D. Puri, R. G. Bhat, A. Surolia, *J. Biol. Chem.* **1993**, *268*, 7668-7677.

[74] D. K. Mandal, L. Bhattacharyya, S. H. Koenig, R. D. I. Brown, S. Oscarson, C. F. Brewer, *Biochemistry* **1994**, *33*, 1157-1162.

[75] M. C. Chervenak, E. J. Toone, *Biochemistry* **1995**, *34*, 5685-5695.

[76] D. Gupta, M. Cho, M. D. Cummings, C. F. Brewer, *Biochemistry* **1996**, *35*, 15236-15243.

[77] A. Surolia, N. Sharon, F. P. Schwarz, *J. Biol. Chem.* **1996**, *271*, 17697-17703.

[78] F. P. Schwarz, H. Ahmed, M. A. Bianchet, L. M. Amzel, G. R. Vasta, *Biochemistry* **1998**, *37*, 5867-5877.

[79] J. V. Pratap, G. M. Bradbrook, G. B. Reddy, A. Surolia, J. Raftery, J. R. Helliwell, M. Vijayan, *Acta Crystallogr.* **2001**, *D57*, 1584-1594.

[80] N. Muller, *Acc. Chem. Res.* **1990**, *23*, 23-28.

[81] H. Beierbeck, L. T. J. Delbaere, M. Vandonselaar, R. U. Lemieux, *Can. J. Chem.* **1994**, *72*, 463-470.

[82] V. R. Srinivas, G. B. Reddy, A. Surolia, *FEBS Lett.* **1999**, *450*, 181-185.

[83] J. E. Leffler, E. Grunwald, *Rates and Equilibria of Organic Reactions*, John Wiley & Sons, New York, **1963**.

[84] G. I. Makhatadze, P. L. Privalov, *J. Mol. Biol.* **1993**, *232*, 639-659.

[85] P. L. Privalov, G. I. Makhatadze, *J. Mol. Biol.* **1993**, *232*, 660-679.

[86] W. Kauzmann, *Adv. Protein Chem.* **1959**, *14*, 1-63.

[87] K. P. Murphy, D. Xie, K. S. Thompson, L. M. Amzel, E. Freire, *Protein Struct. Funct. Genet.* **1993**, *18*, 63-67.

[88] J. M. Sturtevant, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 2236-2240.

[89] R. F. Kelley, M. P. O'Connell, *Biochemistry* **1993**, *32*, 6828-6835.

[90] D. A. Brummell, V. P. Sharma, N. N. Anand, D. Bilous, G. Dubuc, J. Michniewicz, C. R. MacKenzie, J. Sadowska, B. W. Sigurskjold, B. Sinnott, N. M. Young, D. R. Bundle, S. A. Narang, *Biochemistry* **1993**, *32*, 1180-1187.

[91] C. F. Brewer, L. Bhattacharyya, R. D. I. Brown, S. H. Koenig, *Biochem. Biophys. Res. Comm.* **1985**, *127*, 1066-1071.

[92] S. Elgavish, B. Shaanan, *J. Mol. Biol.* **1998**, *277*, 917-932.

[93] T. K. Dam, B. S. Cavada, T. B. Grangeiro, C. F. Santos, V. M. Ceccato, F. A. M. de Susa, S. Oscarson, C. F. Brewer, *J. Biol. Chem.* **2000**, *275*, 16119-16126.

[94] H. C. Siebert, R. Adar, R. Arango, M. Burchert, H. Kaltner, G. Kayser, E. Tajkhorshid, C. W. von der Lieth, R. Kaptein, N. Sharon, J. F. Vliegenthart, H.-J. Gabius, *Eur. J. Biochem.* **1997**, *249*, 27-38.

[95] D. Gupta, T. K. Dam, S. Oscarson, C. F. Brewer, *J. Biol. Chem.* **1997**, *272*, 6388-6392.

[96] N. Navarre, N. Amiot, A. van Oijen, A. Imberty, A. Poveda, J. Jimenez-Barbero, A. Cooper, M. A. Nutley, G.-J. Boons, *Chem. Eur. J.* **1999**, *5*, 2281-2294.

[97] Y. C. Lee, in *Neoglycoconjugates: Preparation and Applications* (Eds.: Y. C. Lee, R. T. Lee), Academic Press, London, **1994**, pp. 3-21.

[98] J. J. Lundquist, E. J. Toone, *Chem. Rev.* **2002**, *102*, 555-578.

[99] W. P. Jencks, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4046-4050.

[100] J. J. Lundquist, S. D. Debenham, E. J. Toone, *J. Org. Chem.* **2000**, *65*, 8245-8250.

[101] S. M. Dimick, S. C. Powell, S. A. McMahon, D. N. Moothoo, J. H. Naismith, E. J. Toone, *J. Am. Chem. Soc.* **1999**, *121*, 10286-10296.

[102] J. B. Corbell, J. J. Lundquist, E. J. Toone, *Tetrahedron: Asymmetry* **2000**, *11*, 95-111.

[103] T. K. Dam, R. Roy, S. K. Das, S. Oscarson, C. F. Brewer, *J. Biol. Chem.* **2000**, *275*, 14223-14230.

[104] T. K. Dam, R. Roy, D. Pagé, C. F. Brewer, *Biochemistry* **2002**, *41*, 1351-1358.

[105] T. K. Dam, R. Roy, D. Pagé, C. F. Brewer, *Biochemistry* **2002**, *41*, 1359-1363.

[106] T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387-429.

[107] Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lönngren, J. Arnarp, M. Haraldsson, H. Lönn, *J. Biol. Chem.* **1983**, *258*, 199-202.

[108] J. Rao, J. Lahiri, L. Isaacs, R. M. Weis, G. M. Whitesides, *Science* **1998**, *280*, 708-711.

[109] J. Rao, L. Yan, J. Lahiri, G. M. Whitesides, R. M. Weis, H. S. Warren, *Chem. Biol.* **1999**, *6*, 353-359.

[110] E. Fan, Z. Zhang, W. E. Minke, Z. Hou, C. L. M. J. Verlinde, W. G. J. Hol, *J. Am. Chem. Soc.* **2000**, *122*, 2663-2664.

[111] R. H. Kramer, J. W. Karpen, *Nature* **1998**, *395*, 710-713.

[112] P. I. Kitov, J. M. Sadowska, G. Mulvey, G. D. Armstrong, H. Ling, N. S. Pannus, R. J. Read, D. R. Bundle, *Nature* **2000**, *403*, 669-672.

[113] H. Ling, A. Boodhoo, B. Hazes, M. D. Cummings, G. D. Armstrong, J. L. Brunton, R. J. Read, *Biochemistry* **1998**, *37*, 1777-1788.

[114] M. Kanai, K. H. Mortell, L. L. Kiessling, *J. Am. Chem. Soc.* **1997**, *119*, 9931-9932.

[115] H. S. Park, Q. Lin, A. D. Hamilton, *J. Am. Chem. Soc.* **1999**, *121*, 8-13.

[116] Y. C. Lee, R. T. Lee, *Neoglycoconjugates: Preparation and Applications*, Academic Press, London, **1994**.

[117] A. C. Roche, P. Midoux, V. Pimpaneau, E. Nègre, R. Mayer, M. Monsigny, *Res. Virol.* **1990**, *141*, 243-249.

[118] R. Roy, D. Pagé, S. F. Perez, V. V. Bencomo, *Glycoconjugate J.* **1998**, *15*, 251-263.

[119] L. L. Kiessling, L. E. Strong, J. E. Gestwicki, *Annu. Rep. Med. Chem.* **2000**, *35*, 321-330.

[120] D. Neumann, O. Kohlbacher, H.-P. Lenhof, C.-M. Lehr, *Eur. J. Biochem.* **2002**, *269*, 1518-1524.

[121] M. A. Clark, B. H. Hirst, M. A. Jepson, *Adv. Drug Deliv. Rev.* **2000**, *43*, 207-223.

[122] A. C. Roche, P. Bailly, M. Monsigny, *INVMDJ* **1985**, *5*, 218-232.

[123] R. Z. Dintzis, M. Okajima, M. H. Middleton, G. Greene, H. M. Dintzis, *J. Immunol.* **1989**, *143*, 1239-1244.

[124] S. A. W. Gruner, E. Locardi, E. Lohof, H. Kessler, *Chem. Rev.* **2002**, *102*, 491-514.

[125] Y. Ohya, H. Oue, K. Nagatomi, T. Ouchi, *Biomacromolecules* **2001**, *2*, 927-933.

[126] E. Lohof, E. Planker, C. Mang, F. Burkhart, M. A. Dechantsreiter, R. Haubner, H.-J. Wester, M. Schwaiger, G. Hölzemann, S. L. Goodman, H. Kessler, *Angew. Chem., Int. Ed. Engl.* **2000**, *39*, 2761-2764.

[127] S. A. W. Gruner, G. Kéri, R. Schwab, A. Venetianer, H. Kessler, *Org. Lett.* **2001**, *3*, 3723-3725.

[128] M. N. Matrosovich, L. S. Mochalova, V. P. Marinina, N. E. Byramova, N. V. Bovin, *FEBS Lett.* **1990**, *272*, 209-212.

[129] H. Kamitakahara, T. Suzuki, N. Nishigori, Y. Suzuki, O. Kanie, C.-H. Wong, *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 1524-1528.

[130] A. Tsuchida, K. Kobayashi, N. Matsubara, T. Muramatsu, T. Suzuki, Y. Suzuki, *Glycoconjugate J.* **1998**, *15*, 1047-1054.

[131] N. K. Sauter, M. D. Bednarski, B. A. Wurzburg, J. E. Hanson, G. M. Whitesides, J. J. Skehel, D. C. Wiley, *Biochemistry* **1989**, *28*, 8388-8396.

[132] G. B. Sigal, M. Mammen, G. Dahmann, G. M. Whitesides, *J. Am. Chem. Soc.* **1996**, *118*, 3789-3800.

[133] S.-K. Choi, M. Mammen, G. M. Whitesides, *J. Am. Chem. Soc.* **1997**, *119*, 4103-4111.

[134] P. M. Simon, P. L. Goode, A. Mobasseri, D. Zopf, *Infect. Immun.* **1997**, *65*, 750-757.

[135] J. D. Klemm, S. L. Schreiber, G. R. Crabtree, *Annu. Rev. Immunol.* **1998**, *16*, 569-592.

[136] D. M. Spencer, T. J. Wandless, S. L. Schreiber, G. R. Crabtree, *Science* **1993**, *262*, 1019-1024.

[137] M. Meldal, P. M. St Hilaire, *Curr. Opin. Chem. Biol.* **1997**, *1*, 552-563.

[138] S. J. Danishefsky, J. R. Allen, *Angew. Chem., Int. Ed. Engl.* **2000**, *39*, 837-863.

[139] G. F. Springer, *Science* **1984**, *224*, 1198-1206.

[140] P. M. St Hilaire, L. Cipolla, A. Franco, U. Tedebark, D. A. Tilly, M. Meldal, *J. Chem. Soc., Perkin Trans. 1* **1999**, 3559-3564.

[141] P. W. Glunz, S. Hintermann, J. B. schwarz, S. D. Kuduk, X.-T. Chen, L. J. Williams, D. Sames, S. J. Danishefsky, V. Kudryashov, K. O. Lloyd, *J. Am. Chem. Soc.* **1999**, *121*, 10636-10637.

[142] L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Curr. Opin. Chem. Biol.* **2000**, *4*, 696-703.

[143] E. J. Gordon, J. E. Gestwicki, L. E. Strong, L. L. Kiessling, *Chem. Biol.* **2000**, *7*, 9-16.

[144] E. J. Gordon, W. J. Sanders, L. L. Kiessling, *Nature* **1998**, *392*, 30-31.

[145] W. J. Sanders, E. J. Gordon, P. J. Beck, R. Alon, L. L. Kiessling, *J. Biol. Chem.* **1999**, *274*, 5271-5278.

[146] S. M. Moghimi, A. R. Rajabi-Siahboomi, *Adv. Drug Deliv. Rev.* **2000**, *41*, 129-133.

[147] N. Hussain, *Adv. Drug Deliv. Rev.* **2000**, *43*, 95-100.

[148] H. Ringsdorf, *J. Polym. Sci.* **1975**, *S51*, 135-153.

[149] V. P. Torchilin, *Eur. J. Pharm. Sci.* **2000**, *11*, S81-S91.

[150] H. Maeda, J. Wu, T. Sawa, Y. Matsumura, K. Hori, *J. Control. Release* **2000**, *65*, 271-284.

[151] V. V. Ranade, *Drug Delivery Syst.* **1989**, *29*, 685-694.

[152] S. Cammas, T. Matsumoto, T. Okano, Y. Sakurai, K. Kataoka, *Mater. Sci. Eng. C* **1997**, *4*, 241-247.

[153] M. Yokoyama, S. Fukushima, R. Uehara, K. Okamoto, K. Kataoka, Y. Sakurai, T. Okano, *J. Control. Release* **1998**, *50*, 79-92.

[154] K. Kataoka, T. Matsumoto, M. Yokoyama, T. Okano, Y. Sakurai, S. Fukushima, K. Okamoto, G. S. Kwon, *J. Control. Release* **2000**, *64*, 143-153.

[155] O. V. Gerasimov, J. A. Boomer, M. M. Qualls, D. H. Thompson, *Adv. Drug Deliv. Rev.* **1999**, *38*, 317-338.

[156] D. Needham, M. W. Dewhirst, *Adv. Drug Deliv. Rev.* **2001**, *53*, 285-305.

[157] N. J. Caplen, E. Kinrade, F. Sorgi, X. Gao, D. Gruenert, D. Geddes, C. Coutelle, L. Huang, E. W. F. W. Alton, R. Williamson, *Gene Ther.* **1995**, *2*, 603-613.

[158] N. K. Egilmez, R. Cuenca, S. J. Yokota, F. Sorgi, R. B. Bankert, *Gene Ther.* **1996**, *3*, 607-614.

[159] M. D. C. Topp, P. J. Dijkstra, H. Talsma, J. Feijen, *Macromolecules* **1997**, *30*, 8518-8520.

[160] J. E. Chung, M. Yokoyama, T. Okano, *J. Control. Release* **2000**, *65*, 93-103.

[161] V. P. Torchilin, *J. Control. Release* **2001**, *73*, 137-172.

[162] G. Molema, D. K. F. Meijer, *Adv. Drug Deliv. Rev.* **1994**, *14*, 25-50.

[163] M. C. Garnett, *Adv. Drug Deliv. Rev.* **2001**, *53*, 171-216.

[164] T. Wileman, C. Harding, P. Stahl, *Biochem. J.* **1985**, *232*, 1-14.

[165] D. K. F. Meijer, P. van der Sluijs, *Pharm. Res.* **1989**, *6*, 105-118.

[166] J. L. Goldstein, M. S. Brown, R. G. W. Anderson, D. W. Russell, W. J. Schneider, *Annu. Rev. Cell Biol.* **1985**, *1*, 1-39.

[167] A. P. Seve, J. Hubert, D. Bouvier, C. Bourgeois, P. Midoux, A. C. Roche, M. Monsigny, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 5997-6001.

[168] L. Fiume, C. Busi, P. Preti, G. Spinosa, *Cancer Drug Deliv.* **1987**, *4*, 145-150.

[169] G. S. Kwon, M. Naito, K. Kataoka, M. Yokoyama, Y. Sakurai, T. Okano, *Colloids Surf. B: Biointerfaces* **1994**, *2*, 429-434.

[170] G. S. Kwon, K. Kataoka, *Adv. Drug Deliv. Rev.* **1995**, *16*, 295-309.

[171] M. Yokoyama, T. Okano, Y. Sakurai, S. Suwa, K. Kataoka, *J. Control. Release* **1996**, *39*, 351-356.

[172] S. A. Hagan, A. G. A. Coombes, M. C. Garnett, S. E. Dunn, M. C. Davies, L. Illum, S. S. Davis, S. E. Harding, S. C. Purkiss, P. R. Gellert, *Langmuir* **1996**, *12*, 2153-2161.

[173] K. Akamatsu, M. Imai, Y. Yamasaki, M. Nishikawa, Y. Takakura, M. Hashida, *J. Drug Target.* **1998**, *6*, 229-239.

[174] M. Hashida, S. Takemura, M. NIshikawa, Y. Takakura, *J. Control. Release* **1998**, *53*, 301-310.

[175] M. J. Lawrence, G. D. Rees, *Adv. Drug Deliv. Rev.* **2000**, *45*, 89-121.

[176] Y. Kaneda, *Adv. Drug Deliv. Rev.* **2000**, *43*, 197-205.

[177] R. Roy, M.-G. Baek, K. Rittenhouse-Olson, *J. Am. Chem. Soc.* **2001**, *123*, 1809-1816.

[178] J. Kreuter, *Adv. Drug Deliv. Rev.* **2001**, *47*, 65-81.

[179] R. L. Juliano, D. Layton, in *Drug Delivery Systems* (Ed.: R. L. Juliano), Oxford University Press, Oxford, **1980**, pp. 189-236.

[180] G. Gregoriadis, *Liposomes as Drug Carriers: Recent Trends and Progress*, John Wiley & Sons, New York, **1988**.

[181] M. Singh, D. O'Hagan, *Adv. Drug Deliv. Rev.* **1998**, *34*, 285-304.

[182] J. Kreuter, *Colloidal Drug Delivery Systems*, Marcel Dekker, New York, **1994**.

[183] S. E. Dunn, A. Brindley, S. S. Davis, M. C. Davies, L. Illum, *Pharm. Res.* **1994**, *11*, 1016-1022.

[184] V. A. Slepushkin, S. Simões, P. Dazin, M. S. Newman, L. S. Guo, M. C. Pedroso de Lima, N. Düzgünes, *J. Biol. Chem.* **1997**, *272*, 2382-2388.

[185] M. Hashida, M. NIshikawa, F. Yamashita, Y. Takakura, *Adv. Drug Deliv. Rev.* **2001**, *52*, 187-196.

[186] G. S. Kwon, T. Okano, *Adv. Drug Deliv. Rev.* **1996**, *21*, 107-116.

[187] G. S. Kwon, *Crit. Rev. Ther. Drug* **1998**, *15*, 481-512.

[188] L. Zhang, A. Eisenberg, *Science* **1995**, *268*, 1728-1731.

[189] P. D. Scholes, A. G. A. Coombes, M. C. Davies, L. Illum, S. S. Davis, in *Controlled Drug Delivery. Challenges and Strategies* (Ed.: K. Park), American Chemical Society, Washington DC, **1997**, pp. 73-106.

[190] T. Riley, S. Stolnik, C. R. Heald, C. D. Xiong, M. C. Garnett, L. Illum, S. S. Davis, S. C. Purkiss, R. J. Barlow, P. R. Gellert, *Langmuir* **2001**, *17*, 3168-3174.

[191] G. S. Kwon, S. Suwa, M. Yokoyama, T. Okano, Y. Sakurai, K. Kataoka, *J. Control. Release* **1994**, *29*, 17-23.

[192] G. Splenhauer, D. Bazile, M. Veillard, C. Prudhomme, J. P. Michalon, *European Patent Office*, 0 520 888 A1, **1992**.

[193] R. Gref, Y. Minamitake, M. T. Peracchia, V. Trubetskoy, V. Torchilin, R. Langer, *Science* **1994**, *263*, 1600-1603.

[194] I. L. Shin, S. Y. Kim, Y. M. Lee, C. S. Cho, Y. K. Sung, *J. Control. Release* **1998**, *51*, 1-11.

[195] T. Govender, S. Stolnik, C. D. Xiong, S. Zhang, L. Illum, S. S. Davis, *J. Control. Release* **2001**, *75*, 249-258.

[196] G. S. Kwon, M. Naito, M. Yokoyama, T. Okano, Y. Sakurai, K. Kataoka, *Pharm. Res.* **1995**, *12*, 192-195.

[197] K. Kataoka, A. Harada, Y. Nagasaki, *Adv. Drug Deliv. Rev.* **2001**, *47*, 113-131.

[198] K. Kataoka, T. Matsumoto, M. Yokoyama, T. Okano, Y. Sakurai, S. Fukushima, K. Okamoto, G. S. Kwon, *J. Control. Release* **2000**, *64*, 143-153.

[199] K. Yamada, M. Minoda, T. Miyamoto, *Macromolecules* **1999**, *32*, 3553-3558.

[200] K. Yasugi, T. Nakamura, Y. Nagasaki, M. Kato, K. Kataoka, *Macromolecules* **1999**, *32*, 8024-8032.

[201] K. Kataoka, G. S. Kwon, M. Yokoyama, T. Okano, Y. Sakurai, *J. Control. Release* **1993**, *24*, 119-132.

[202] Y. Kakizawa, A. Harada, K. Kataoka, *Biomacromolecules* **2001**, *2*, 491-497.

[203] A. Rösler, G. W. M. Vandermeulen, H.-A. Klok, *Adv. Drug Deliv. Rev.* **2001**, *53*, 95-108.

[204] H. Huang, E. E. Remsen, T. Kowalewski, K. L. Wooley, *J. Am. Chem. Soc.* **1999**, *121*, 3805-3806.

[205] T. Sanji, Y. Nakatsuka, S. Ohnishi, H. Sakurai, *Macromolecules* **2000**, *33*, 8524-8526.

[206] J. Liu, K. L. Wooley, *Polym. Mater. Sci. Eng.* **2001**, *84*, 967-968.

[207] K. S. Murthy, Q. Ma, C. G. J. Clark, E. E. Remsen, K. L. Wooley, *J. Chem. Soc., Chem. Commun.* **2001**, 773-774.

2

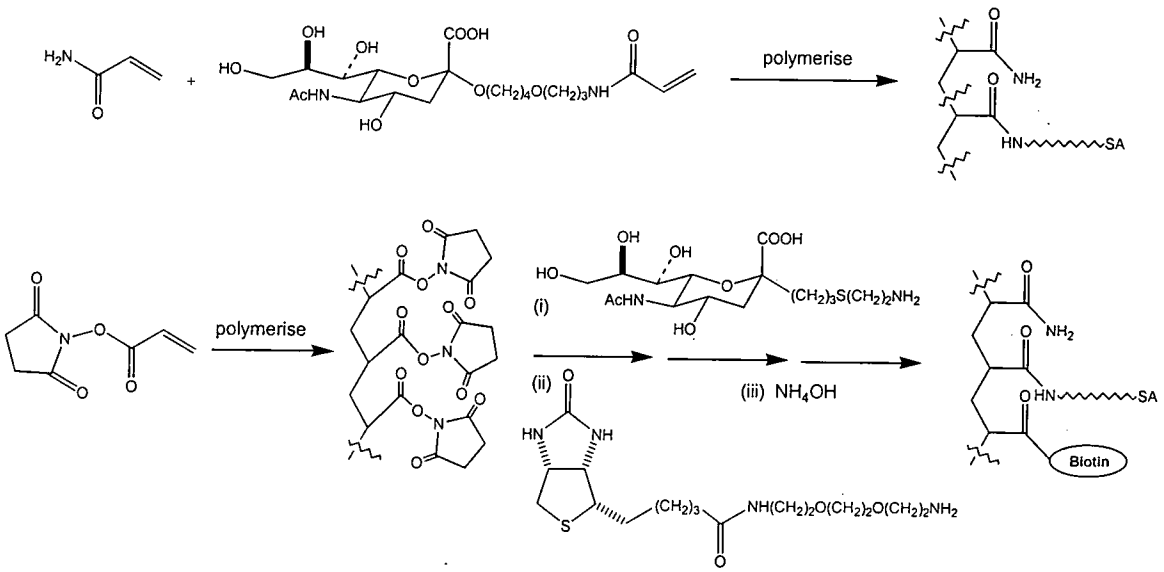
GLYCOSYLATED POLYMETHACRYLATES: SYNTHESIS AND CHARACTERISATION

2.1 INTRODUCTION

Polymers consisting of a chemically and biologically stable C-C backbone and a hydrophilic saccharide moiety in the side chain are called "glycopolymers"[1]. Due to their application in basic biochemical and biomedical research, glycopolymers have attracted increasing attention since they were first developed by Horejsi *et al.* in 1978[2].

The synthesis of carbohydrate-based polymers usually requires the preparation of polymerisable sugar derivatives[3-15], although the less frequently employed glycosylation of polymers is also possible[16,17]. Given that the synthesis of glycosylated monomers often requires multi-step reactions, it is somewhat surprising that post-polymerisation glycosylation of already defined polymers has not been more widely applied. Whitesides and co-workers[18] have compared pre- and post-polymerisation glycosylation techniques in the preparation of polyacrylamides carrying sialic acid residues (Scheme 1). They demonstrated that post-polymerisation glycosylation resulted in glycopolymers that were more effective inhibitors (*ca.* 100-fold) of the hemagglutination of erythrocytes by influenza virus. The authors hypothesised that the observed differences in inhibitory activity might be related to differences in the distribution of sialic acid groups along the polymer backbone. They suggested that polymers obtained by copolymerisation of acrylamide and the acrylamide derivative bearing the carbohydrate moiety might not have a random

distribution of the glycosyl residues, due to differences in the polymerisation rates of the two monomers. The uneven distribution of sialic acid groups limits the surface of the virus that can interact with the polymer, decreasing the binding constant. In contrast, post-polymerisation glycosylation is more likely to give a random distribution of sugar moieties, since the latter parameter is in this case controlled by differences in the reactivity of the side chains of poly(*N*-acryloyloxysuccinimide). These differences are probably small because neutral polyacrylates in good solvents tend to behave as random coils. The introduction of biotin into the polymers allowed ELISA assays, which demonstrated that the inhibition of agglutination correlated well with the affinity of the polymers for the lectin.



Scheme 1 Pre- (top) and post-polymerisation (bottom) glycosylation techniques.

Perhaps the reason why post-polymerisation glycosylation procedures have not been more widely employed has to be sought for in the lack of methods for high yielding well-defined glycosylation of macromolecules[19]. Glycosylation of preformed polymers is usually incomplete, due to steric hindrance. Thus, the glycopolymers obtained present an ill-defined structure, with some functional groups

still unreacted. To carry out reactions as completely as possible, less reactive pendant functional groups have to be appropriately activated before the glycosylation step[20]. In general, glycopolymers synthesised by glycosylation of preformed polymers are more suitable for practical purposes where a strictly well-defined structure is not required.

Different types of glycopolymers have been synthesised, mainly polyacrylamides, polystyrenes, polyacrylates and polymethacrylates. *N*-(2-hydroxypropyl)methacrylamide, HMPA, copolymers containing fucosylamine and 5-aminosalicylic acid (5-ASA, a drug for the treatment of ulcerative colitis) have been prepared by Kopeček and co-workers[21,22]. The materials have been shown to be suitable as colon-specific drug delivery systems. Whilst the fucosylamine allows the targeting by interaction with specific colonic mucosal lectins, the drug, bound through aromatic azo bonds, can be released by the degradative action of microbial azoreductases present only in the colon. Kobayashi *et al.*[23] have synthesised a variety of carbohydrate-containing polystyrene derivatives, which have been used as cell-specific biomedical materials. In particular, lactose-carrying polystyrene (PVLA) has been shown to be a useful *substratum* for the culture of hepatocytes. The adhesion of hepatocytes is attained by the specific molecular recognition between the hepatic asialoglycoprotein receptors and the highly concentrated galactose residues on the polymer chains. Polystyrene derivatives bearing sialyllactose moieties have been prepared by Kobayashi and co-workers[24] from a mixture of (α-2,3) and (α-2,6) linked sialyllactose isomers isolated from bovine milk. The reducing ends of the oligosaccharides were converted to amino groups and then coupled with *p*-vinylbenzoyl chloride. The obtained styrene derivatives were polymerised by free radical polymerisation and the interaction of the resulting glycopolymers with

influenza virus was investigated. The glycopolymers inhibited the hemagglutination of influenza A virus 1000-fold stronger than the corresponding oligosaccharides themselves. Moreover, various viruses could be trapped by the glycopolymers adsorbed on a polystyrene surface. The inhibitory and trapping affinities correlated with the sialyl linkage specificities of the virus strains. Ohno *et al.*[25] reported the synthesis of PVLA obtained by means of a different route.

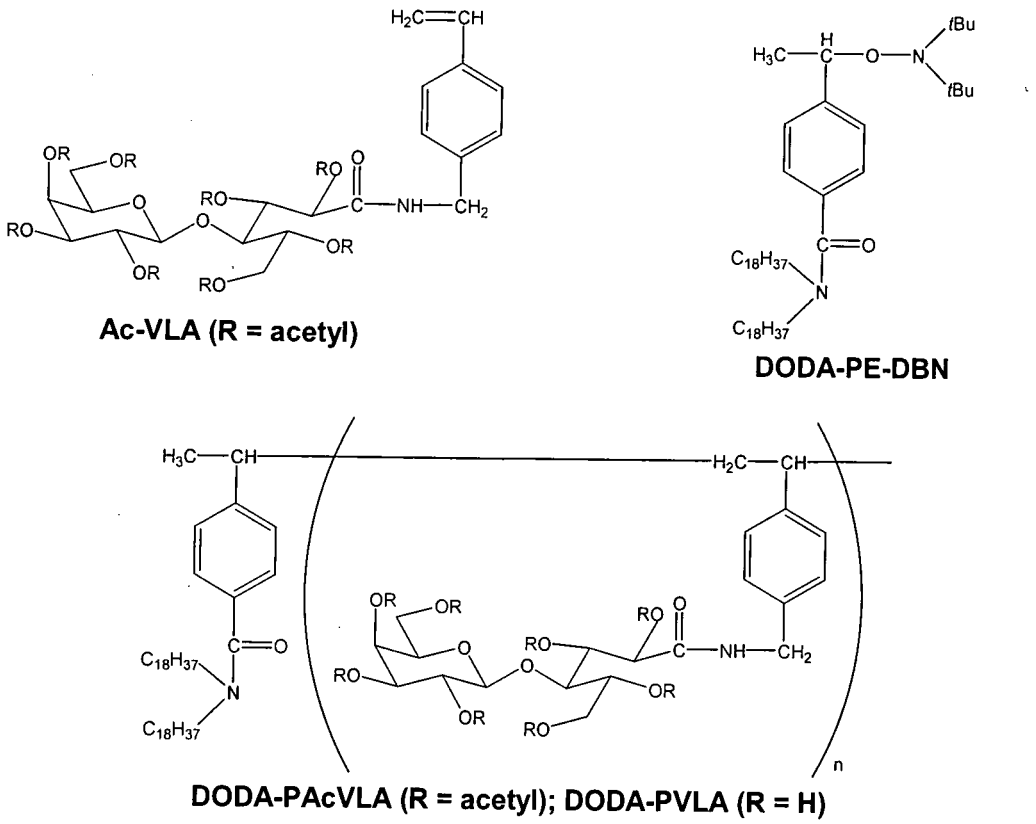


Figure 1 Structures of Ac-VLA, DODA-PE-DBN, DODA-PAcVLA and DODA-PVLA.

The protected glycosylated monomer was polymerised by nitroxide-mediated free radical polymerisation with a lipophilic alkoxyamine "initiator" (Figure 1). The deprotected polymer DODA-PVLA was then used to prepare sugar-carrying liposomes, the galactose residues of which were specifically and effectively recognised by the galactose-specific lectin RCA. Polyacrylamides containing glucose

and galactose units were synthesised from poly(acryloyl chloride) using glucosamine and galactosamine, respectively[26]. The attachment of different kinds of cells to tissue culture polystyrene plates coated with the glycopolymers was then investigated. Glycopolymers bearing triaantennary *N*-acetyllactosamine were prepared by radical copolymerisation of the glycomonomer and acrylamide. Enzymatic sialylation of the copolymer resulted in glycopolymers bearing different sialooligosaccharides depending on the enzyme used[27] (Figure 2).

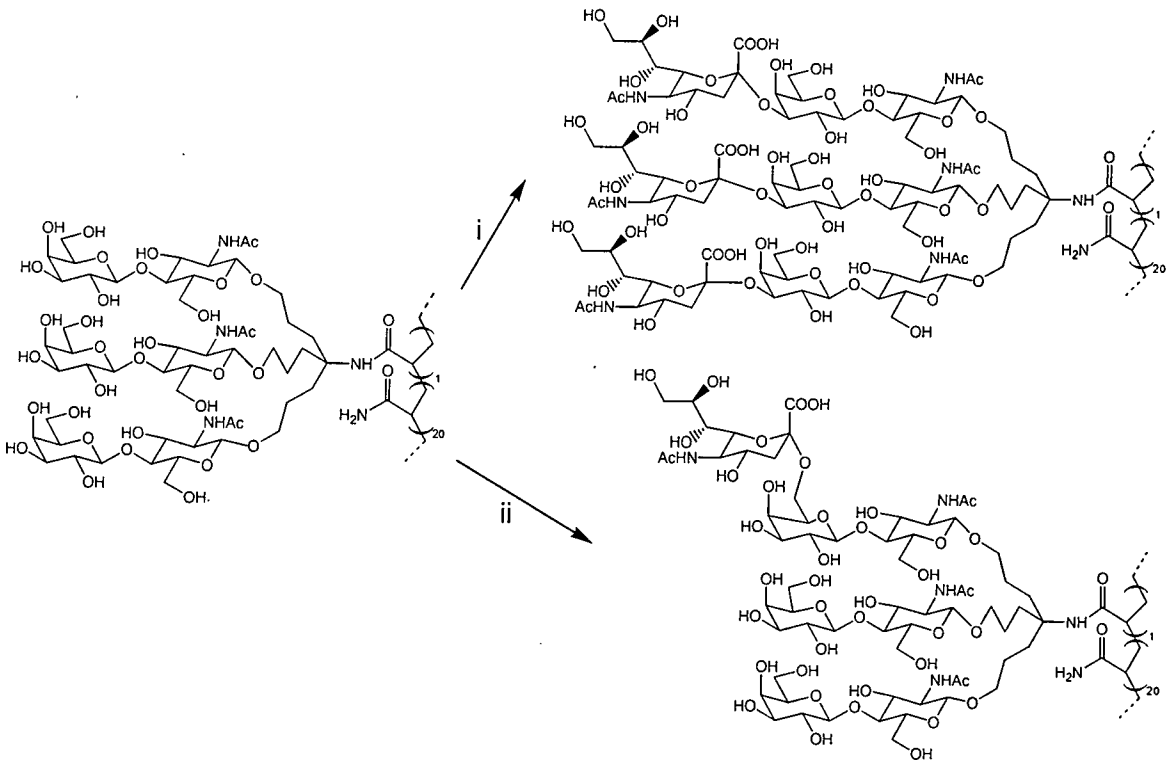
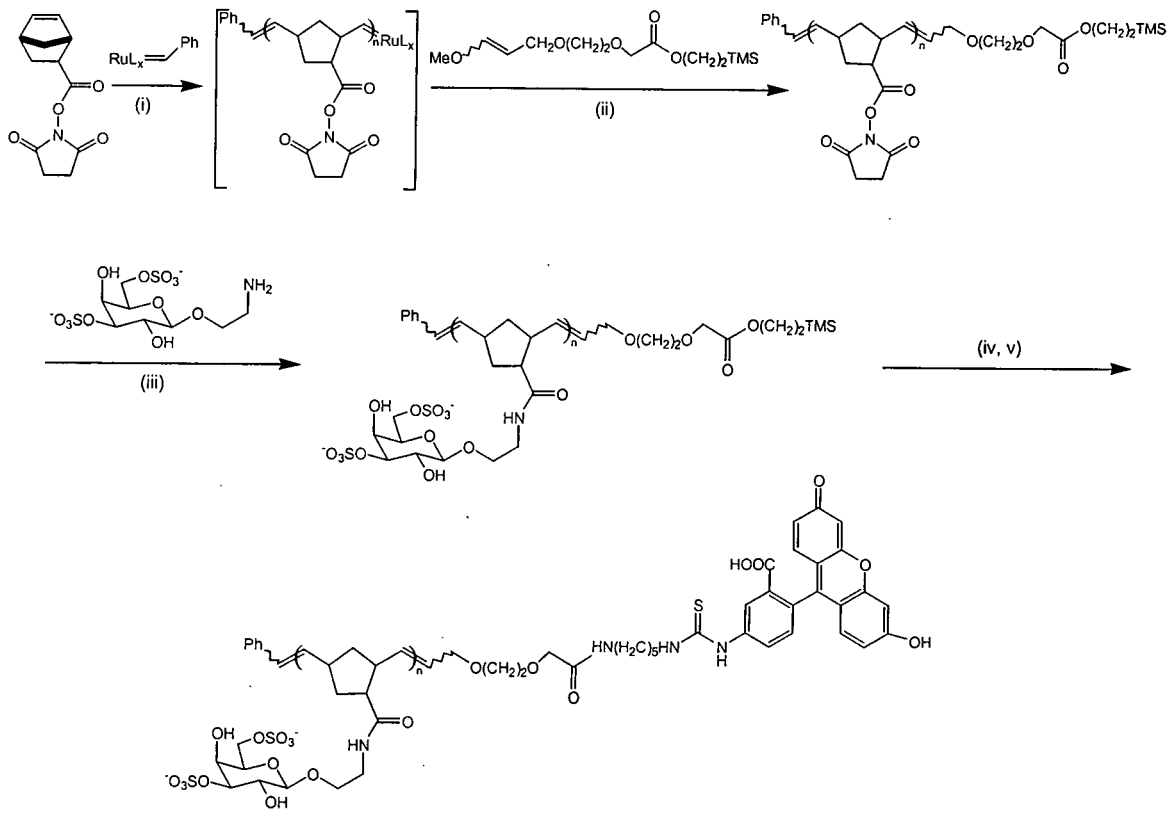


Figure 2 Structures of polymers carrying sialooligosaccharides. (i), α-2,3-sialyltransferase from porcine liver; (ii), α-2,6-sialyltransferase from rat liver.

Poly(*N*-*p*-vinylbenzyl-D-glucoronamide) (PV6Gna) was synthesised by free radical polymerisation of the corresponding monomer[28]. Surprisingly, hepatocytes were shown to adhere to polystyrene dishes coated with the glycopolymer. It was hypothesised that the mechanism of the interaction between PV6Gna and the cells is different from that previously mentioned for PVLA and probably mediated by a

different C-type lectin. Ring-opening metathesis polymerisation (ROMP) was used by Kiessling *et al.*[29] for the preparation of several glycopolymers, including that reported in Scheme 2. The attachment of the fluorescein derivative allowed the visualisation of the polymer binding to cells displaying L-selectins. The results suggested that the polymer binds specifically to cell surface L-selectins through multivalent interactions.



Scheme 2 Synthesis of glycopolymers by ROMP. (i), dichloroethane, DCE, 30 min; (ii), excess of end-capping agent; (iii), diisopropylcarbodiimide, DIC; (iv), 50 mM NaOH, 60 °C, 2 h; (v), 5-((5-aminopentyl)-thioureidyl) fluorescein, EDCl, *N*-hydroxysulfosuccinimide, H_2O, 24 h.

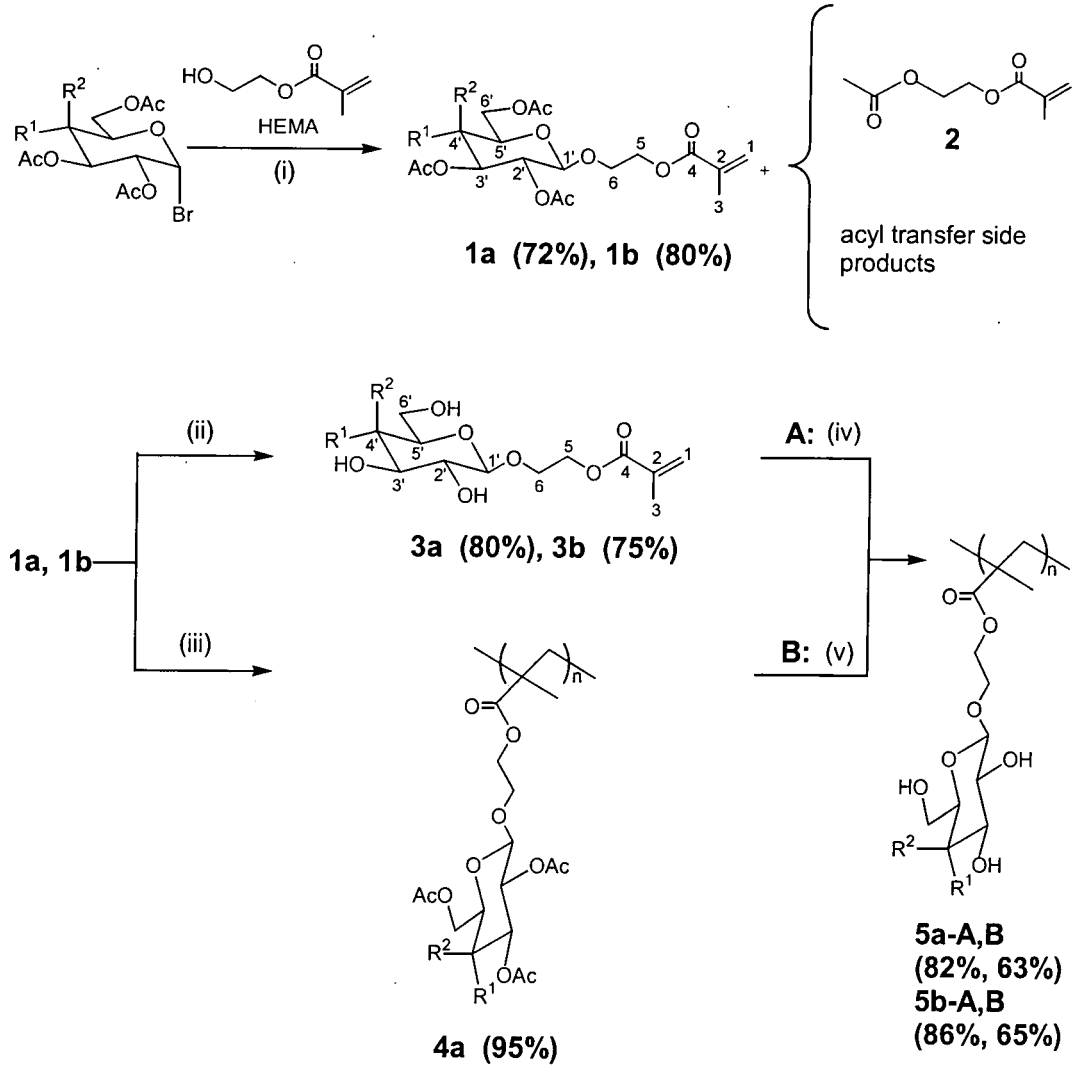
Baek and Roy[30] reported the synthesis of a series of T-antigen containing glycopolymers. The materials were prepared from poly[*N*-(acryloxy)succinimide] by amidation with an amine-ending T-antigen derivative and various amines of increasing alkyl chain length. The glycopolymers were used as coating antigens for

solid-phase ELISA. It was shown that the most lipophilic polymer exhibited the strongest antigen-antibody interaction.

Tailor-made glycopolymers were prepared by Nagahori and Nishimura[31]. A polymerisable glycomonomer, containing a trimannoside residue, was incubated with concanavalin A (template lectin as guest molecule) in a buffer solution and copolymerised with acrylamide and a cross-linker. The obtained cross-linked polymer showed higher affinity for the lectin compared to the corresponding linear glycopolymer, indicating that specific cavities for the guest proteins were constructed in the glycoligand by the adopted template-based polymerisation strategy.

Due to the polyfunctionality of sugars, multistep reactions including protecting-group chemistry are typically required for their manipulation. Indeed, selective reactions of non-protected sugars to form polymerisable monomers have usually been achieved by enzymatic catalysis[32,33]. The removal of protecting groups can be carried out either before or after the polymerisation. When considering the point at which to deprotect, the lability of the monomer and, to a lesser extent, the polymer, together with the possibility of a non-quantitative deprotection of the polymer, all have to be taken in account. Usually deprotection of the polymers is not quantitative; therefore the removal of protecting groups is, if possible, preferably carried out at the monomer stage[8]. Nevertheless, to our knowledge, only few examples of deprotection at the monomer stage have been reported so far[4,6]. Despite the importance of this strategy, no systematic comparison has been made between deprotection pre- or post-polymerisation.

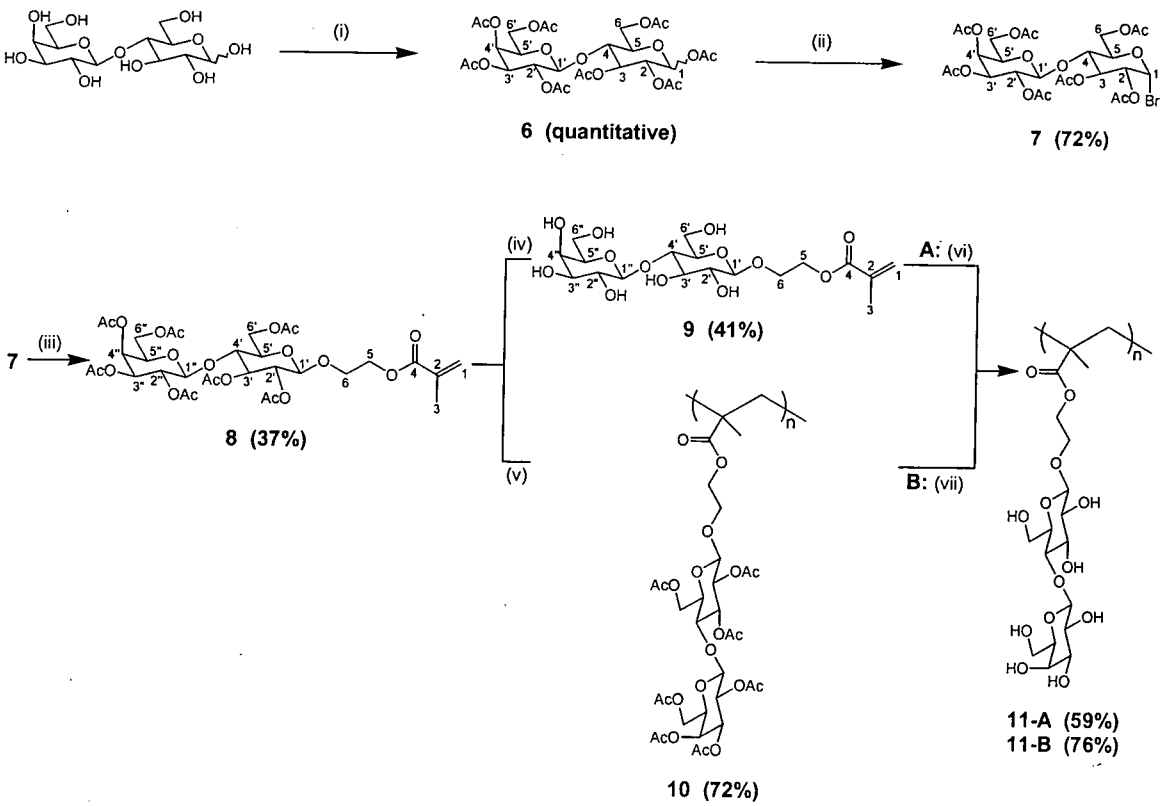
In this study methacrylate derivatives of glucose and galactose were synthesised by glycosylation of 2-hydroxyethyl methacrylate (HEMA) following the procedure reported in Schemes 3 and 4[34].



1a, 4a R^1 = OAc R^2 = H **1b, 4b** R^1 = H R^2 = OAc

3a, 5a R^1 = OH R^2 = H **3b, 5b** R^1 = H R^2 = OH

Scheme 3 Synthesis of monomers and polymers carrying glucose and galactose residues. *Reagents and conditions*: (i), HEMA (3 eq.), AgOTf (1.2 eq.), dry DCM, under N_2, -40 °C, 48 h; (ii), $NaOCH_3$ [cat.], dry CH_3OH, RT, under N_2; (iii), AIBN, $CHCl_3$, 65 °C, 48 h; (iv), $K_2S_2O_8$, H_2O:CH_3OH (4:1), 65 °C, 48 h; (v), $NaOCH_3$ [cat.], CH_3OH:$CHCl_3$ (1:1), RT, under N_2.



Scheme 4 Synthesis of monomer and polymer carrying lactose residues. *Reagents and conditions*: (i), Ac_2O/NaOAc, reflux; (ii), HBr/CH_3COOH (30%), dry DCM; (iii), HEMA (3 eq.), AgOTf (1.2 eq.), dry DCM, under N_2, -40 °C, 48 h; (iv), $NaOCH_3$ [cat.], dry CH_3OH, RT, under N_2; (v), AIBN, $CHCl_3$, 65 °C, 48 h; (vi), $K_2S_2O_8$, H_2O:CH_3OH (4:1), 65 °C, 48 h; (vii), $NaOCH_3$ [cat.], CH_3OH:$CHCl_3$ (1:1), RT, under N_2.

The protected monomers were first deacetylated with sodium methoxide in methanol and then polymerised in a mixture 4:1 of water and methanol at 65 °C using potassium persulfate as initiator (route A).Alternatively, the protected monomers were polymerised with 2,2'-azoisobutyronitrile (AIBN) in chloroform at 65 °C and the resulting polymers were subsequently deprotected with sodium methoxide in a 1:1 mixture of chloroform and methanol (route B).

Methacrylate and acrylate derivatives containing glucose and galactose have been synthesised previously. Kitazawa *et al.*[35] reported the glycosylation of HEMA and HEA using several methyl glycosides as glycosyl donors, including methyl glucoside and methyl galactoside, in the presence of phosphomolybdic acid as catalyst

and 2,4-dinitrochlorobenzene as a polymerisation inhibitor. However, the stereoselectivity of this method is low and a ratio of α,β anomers is obtained. Nakaya *et al.*[36] synthesised 2-(2',3',4',6'-tetra-*O*-acetyl-β-D-glucopyranosyloxy)ethyl methacrylate by reacting HEMA with 2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosylbromide by the method of Helferich[37], in the presence of silver oxide or mercury(II) cyanide, with yields of 54% and 58% respectively. After free radical polymerisation, the polymer obtained was deacetylated with sodium methoxide and the title polymer identified by infrared spectroscopy data alone. 2-(2',3',4',6'-tetra-*O*-acetyl-β-D-glucopyranosyloxy)ethyl acrylate (AcGEA) was synthesised by Liang *et al.*[38-40] according to the method of Helferich[37], by glycosylation of HEA with 2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosylbromide using mercury(II) bromide as the catalyst. After polymerisation by conventional free radical and atom transfer radical polymerisation, the polymer was deprotected with sodium methoxide. The self-association tendency of PGEA in water was studied; a dependence of the critical aggregation concentration on the molecular weight of the polymer and on the temperature was observed. In order to understand the influence of hydrophobicity on the critical aggregation concentration, AcGEA was copolymerised with stearyl acrylate in different ratios. β-D-galactopyranosyloxyethyl methacrylate has so far been synthesised in a stereo-controlled manner exclusively by enzymatic catalysis[32,33].

In light of these disparate routes, some with low yields or stereoselectivities and/or involving toxic mercury salts, we derived parallel routes to protected **1a**,**b**, **8** and deprotected **3a**,**b**, **9**. Not only was β-stereoselectivity in the glycosylation reactions ensured by neighbouring group participation of acetyl groups at O-2 in the glycosyl donors, but also the yields of these monomer syntheses were considerably

improved compared to previous syntheses. It is worthwhile to note that stereoselective synthetic procedures represent an essential requirement for the preparation of glycoconjugates with potential applications in medicinal chemistry. Only one anomer is, in fact, often preferentially recognised by the specific biological receptor. Thus, for example, the asialoglycoprotein receptor on the surface of hepatocytes binds efficiently only to β-galactosyl residues. Furthermore, also when the receptor lacks high anomeric selectivity, its affinity toward the two stereoisomers may still be different. Therefore, if, for example, the binding of a glycopolymer to a certain specific protein is under investigation, different sugar units' compositions would lead to significantly different results (see Chapter 4). All monomeric and polymeric products have been characterised fully by infrared and NMR spectroscopy, optical activity measurements, elemental analysis and, for the monomers, mass spectrometry, leading to a more thorough characterisation than those reported so far[35,36,38-40] – critical for a proper comparison of the influence of sugar type in biological applications. Absolute molecular weights were obtained for all polymers by size exclusion chromatography (SEC) using the previously estimated refractive index increments, dn/dc. Good characterisation, high purity, and a fully known composition are important requirements for the precise use and application of glycopolymers and for a true understanding of any results obtained. Therefore, in this study a detailed characterisation of the monomers and polymers was carried out and for the first time we investigated how the polymer composition depends on the method of preparation used.

2.2 RESULTS AND DISCUSSION

2.2.1 GLYCOMONOMERS

The protected monomers AcGlcEMA (**1a**), AcGalEMA (**1b**) and AcLacEMA (**8**) were synthesised by coupling the glycosyl donors, 2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl bromide, 2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranosyl bromide and 2',3',4',6'-tetra-*O*-acetyl-β-D-galactopyranosyl (1'→4)-2,3,6-tri-*O*-acetyl-α-D-glucopyranosyl bromide (**7**), respectively, with the glycosyl acceptor 2-hydroxyethyl methacrylate (HEMA), in dry dichloromethane, using silver trifluoromethanesulfonate as a catalyst. For the disaccharide, lactose needed to be converted to the more reactive peracetylated bromide, carrying a good leaving group at the anomeric centre, before the coupling reaction. The synthetic procedures are reported in Schemes 3 and 4, for **1a,b** and **8**, respectively. β-stereoselectivity in the glycosylation reactions was ensured by neighbouring group participation of acetyl groups at O-2 in the glycosyl donors. Acetyl transfer to the nucleophilic alcohol (HEMA) led to the formation of AcEMA (**2**) as the major side product. Such side reactions have been reported previously[16,41-43] and may be attributed to rearrangements of ortho ester intermediates[16]. According to several proposals[44,45] in the literature, the glycosylation reaction proceeds by activation of the glycosyl donor by silver triflouromethanesulfonate, leading to the irreversible formation of a glycosyl oxocarbenium ion[46] that, due to neighbouring group participation, is in equilibrium with the corresponding carbocationic species[16,47]. Nucleophilic attack of HEMA on the latter species can then result in the formation of the desired products **1a**, **1b** and **8** with AcEMA and monodeacetylated compounds as side products. Intramolecular neighbouring group participation is

expected to be kinetically favoured with respect to intermolecular nucleophilic attack so that the oxocarbenium ion is unlikely to have a long lifetime[48,49].

Products **1a**, **1b**, **2** and **8** were isolated by flash column chromatography. 2-(*O*-acetyl)ethyl methacrylate (AcEMA, **2**) was characterised by NMR spectroscopy and mass spectrometry. In addition to **2**, unreacted HEMA was also present in each crude product; this was not separable from the desired product by column chromatography. This problem was overcome by acetylation of HEMA at the end of the glycosylation reaction to give **2**, which was removed easily by flash chromatography. In this manner, AcGlcEMA, AcGalEMA and AcLacEMA were obtained in yields of 72%, 80% and 37%. It is worth noting that the yields for products **1a** and **1b** are significantly higher than those previously reported. All compounds were characterised by IR and NMR spectroscopy, optical activity measurements, mass spectrometry and elemental analysis. All glycomonomers were identified by ^{1}H-NMR spectroscopy as being the β-anomers.

AcGlcEMA was obtained as a colourless crystalline solid* and its crystal structure was determined by single-crystal X-ray diffraction (Figure 3). The heterocycle adopts a normal chair conformation, with all the substituents in equatorial orientations. The olefinic C(10)=C(12) bond and the adjacent ester group are nearly coplanar: the O(9)C(9)C(10)C(12) torsion angle is 3.1(3)°. AcGalEMA was obtained as colourless oil that was stored at 4 °C after addition of a radical polymerisation inhibitor (hydroquinone, 10 ppm). AcLacEMA was obtained as an amorphous solid after freeze-drying.

The glycomonomers are soluble in chloroform, methanol, dichloromethane, tetrahydrofuran, acetone, benzene and DMF.

* Crystal data are reported in Appendix I.

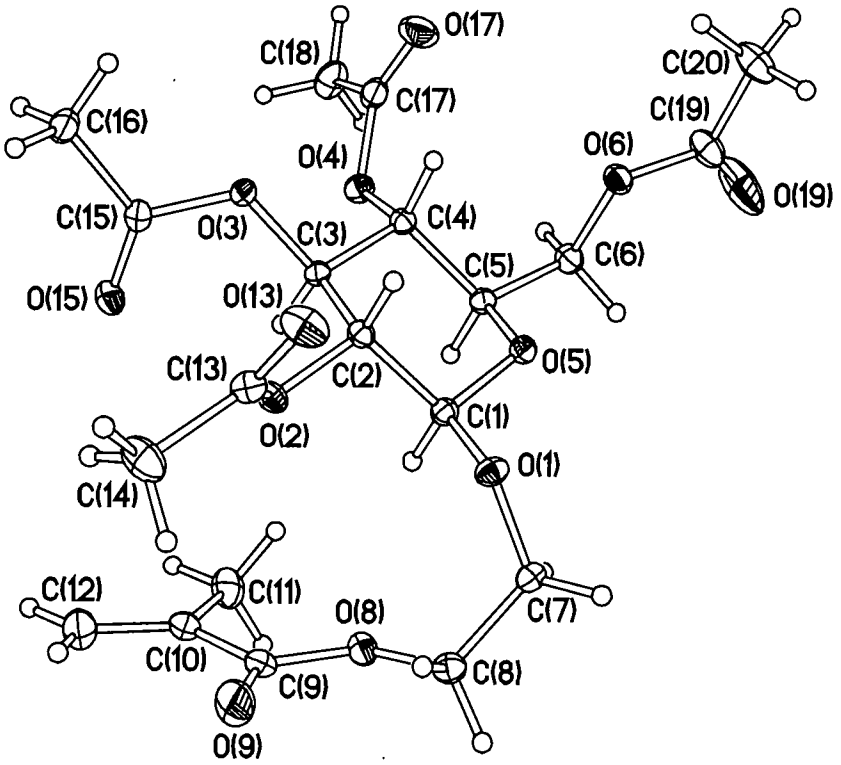


Figure 3 Crystal structure of AcGlcEMA (**1a**).

The monomers were deprotected using a catalytic quantity of sodium methoxide in methanol, as reported in Schemes 3 and 4. The reaction, monitored continuously by TLC (acetonitrile:water, 9:1), was stopped when the product resulting from the cleavage of the ester bond of the HEMA moiety was observed (as a spot having R_f = 0.2 for the monosaccharide derivatives and R_f = 0.1 for the monomer carrying the disaccharide residue). The purification of GlcEMA (**3a**) and GalEMA (**3b**) by chromatography (chloroform:methanol, 8:2) afforded the products in high purity. GlcEMA and GalEMA were obtained as strongly hygroscopic amorphous colourless solids, with yields of 80% and 75%, respectively. LacEMA (**9**) was isolated by chromatography (dichloromethane:methanol:acetic acid:water, 70:20:1:2). After removal of solvent under vacuum and freeze-drying, the product was obtained as an amorphous solid, with yield of 41%.

All deprotected monomers were fully characterised for the first time by IR and NMR spectroscopy, optical activity measurements, mass spectrometry and elemental analysis (Table 1). ES-mass spectra showed only one signal each, corresponding to the completely deprotected products. Infrared spectra showed the complete disappearance of the carbonyl absorption bands at *ca.* 1750 cm^{-1} corresponding to the *O*-acetyl protecting groups. No signals due to the protecting groups could be seen in the ^{1}H- and ^{13}C-NMR spectra.

GlcEMA, GalEMA and LacEMA are soluble in methanol, DMF and water.

Table 1 Yields and properties of monomers.

Monomer	Yield (%)	$[\alpha]_D^T$	Elemental Analysis
AcGlcEMA (1a)	72	-12.8 (*c* 0.14, $CHCl_3$, 28 °C)	**$C_{20}H_{28}O_{12}$ (460.43)** calc. C = 52.17%, H = 6.13% found C = 52.11%, H = 6.19%
AcGalEMA (1b)	80	-5.5 (*c* 0.25, $CHCl_3$, 28 °C)	**$C_{20}H_{28}O_{12}$ (460.43)** calc. C = 52.17%, H = 6.13% found C = 51.94%, H = 6.03%
AcLacEMA (8)	37	-16.9 (*c* 0.10, $CHCl_3$, 23 °C)	**[$C_{32}H_{44}O_{20}$ (748.68) + 0.25 mol H_2O]** calc. C = 51.03%, H = 5.96 % found C = 50.82%, H = 5.96%
GlcEMA (3a)	80	-22.5 (*c* 0.20, CH_3OH, 20 °C)	**$C_{12}H_{20}O_8$ (292.28)** calc. C = 49.31%, H = 6.90% found C = 48.99%, H = 7.01%
GalEMA (3b)	75	-5.5 (*c* 0.20, CH_3OH, 20 °C)	**[$C_{12}H_{20}O_8$ (292.28) + 0.3 mol H_2O]** calc. C = 48.42%, H = 6.97 % found C = 48.34%, H = 6.96%
LacEMA (9)	41	-7.6 (*c* 0.10, H_2O, 24 °C)	**$C_{18}H_{30}O_{13}$ (454.42)** calc. C = 47.58%, H = 6.65% found C = 47.31%, H = 6.46%

2.2.2 GLYCOPOLYMERS

pAcGlcEMA (**4a**), pAcGalEMA (**4b**) and pAcLacEMA (**10**) were obtained as white solids by polymerising **1a**, **1b** and **8**, respectively, with 2,2'-azoisobutyronitrile (AIBN) in chloroform at 65 °C for 48 h. The IR spectra showed that the vinyl absorption bands (1321 and 1299 cm^{-1}, 1320 and 1295 cm^{-1}, 1321 and 1298 cm^{-1} for AcGlcEMA, AcGalEMA and AcLacEMA, respectively) had disappeared. Moreover, no signals due to vinyl protons and carbons could be seen in the ^{1}H- and ^{13}C-NMR spectra.

Table 2 Yields and properties of polymers.

Polymer	Yield (%)	$[\alpha]_D^T$	Elemental Analysis
pAcGlcEMA (4a)	95	-11.7 (*c* 0.11, $CHCl_3$, 22 °C)	**$(C_{20}H_{28}O_{12})_m$, $(460.43)_m$** calc. C = 52.17%, H = 6.13% found C = 52.03%, H = 6.17%
pAcGalEMA (4b)	93	-12.2 (*c* 0.11, $CHCl_3$, 22 °C)	**$(C_{20}H_{28}O_{12})_m$, $(460.43)_m$** calc. C = 52.17%, H = 6.13% found C = 51.92%, H = 6.26%
pAcLacEMA (10)	72	–31.0 (*c* 0.12, $CHCl_3$, 23 °C)	**[$(C_{32}H_{44}O_{20})_m$, $(748.68)_m$ + 0.5 mol H_2O]** calc. C = 50.73%, H = 5.99% found C = 50.71%, H = 5.90%
pGlcEMA (5a-A)	82	- 15.2 (*c* 0.12, H_2O, 22 °C)	**[$(C_{12}H_{20}O_8)_m$, $(292.28)_m$ + 0.1 mol H_2O]** calc. C = 49.01%, H = 6.92% found C = 48.82%, H = 6.92%
pGalEMA (5b-A)	86	+ 8.0 (*c* 0.10, H_2O, 22 °C)	**[$(C_{12}H_{20}O_8)_m$, $(292.28)_m$ + 0.1 mol H_2O]** calc. C = 49.01%, H = 6.92% found C = 48.84%, H = 6.91%
pLacEMA (11-A)	59	-12.7 (*c* 0.11, H_2O, 24 °C)	**[$(C_{18}H_{30}O_{13})_m$, $(454.42)_m$ + 0.3 mol H_2O]** calc. C = 47.02%, H = 6.71% found C = 46.79%, H = 6.82%
pGlcEMA (5a-B)	63	- 10.0 (*c* 0.08, H_2O, 22 °C)	**$(C_{12}H_{20}O_8)_m$, $(292.28)_m$** calc. C = 49.31%, H = 6.90% *found C = 51.26%, H = 6.69%*
pGalEMA (5b-B)	65	+ 15.7 (*c* 0.09, H_2O, 22 °C)	**$(C_{12}H_{20}O_8)_m$, $(292.28)_m$** calc. C = 49.31%, H = 6.90% *found C = 51.20%, H = 6.75%*
pLacEMA (11-B)	76	-6.5 (*c* 0.10, H_2O, 24 °C)	**$(C_{18}H_{30}O_{13})_m$, $(454.42)_m$** calc. C = 47.58%, H = 6.65% *found C = 51.31%, H = 6.38%*

As for the corresponding monomers, pAcGlcEMA, pAcGalEMA and pAcLacEMA showed optical activity due to the saccharide units. The specific rotations $[\alpha]_D^T$, measured in chloroform, are reported in Table 2.

All protected glycopolymers are soluble in chloroform, THF, benzene, DMF and acetone.

The number- and weight-average molecular weights were determined by size exclusion chromatography (Table 3).

Table 3 Absolute molecular weight as determined by SEC and dn/dc of polymers.

Polymer	M_n g mol^{-1}	M_w g mol^{-1}	M_w/M_n	dn/dc ml g^{-1}
pAcGlcEMA (4a)	61,600	135,900	2.21	0.058 ± 0.001
pAcGalEMA (4b)	63,000	156,870	2.49	0.071 ± 0.002
pAcLacEMA (10)	35,100	43,500	1.24	0.060 ± 0.002
pGlcEMA (5a-A)	23,000	60,000	2.61	0.131
pGalEMA (5b-A)	461,000	1,019,000	2.21	0.140
pLacEMA (11-A)	296,000	669,200	2.26	0.150
pGlcEMA (5a-B)	5,932	10,410	1.75	0.126
pGalEMA (5b-B)	2,984	6,410	2.15	0.146
pLacEMA (11-B)	15,380	21,840	1.42	0.142

4a, **4b** and **10**: SEC performed in THF; dn/dc, determined in THF at 30 °C using a calibrated light scattering signal, working at known concentrations and with a He-Ne laser (wavelength 670 nm) as the light source.
5a-A,B, **5b-A,B** and **11-A,B**: SEC in aqueous solution; dn/dc determined by means of a differential refractometer, at 25 °C, working in the same experimental conditions used for the SEC measurements.

Absolute molecular weight values were obtained using the previously estimated refractive index increments, dn/dc, which were determined in THF at 30 °C using a calibrated light scattering signal, working at known concentrations and with a He-Ne laser (wavelength 670 nm) as the light source. Values of 0.058 ml g^{-1}, 0.071 ml g^{-1}

and 0.060 ml g^{-1} for pAcGlcEMA, pAcGalEMA and pAcLacEMA, respectively, were obtained. These values were consistent for three batches of the same polymer. Since the experimental conditions were the same for all polymer solutions, this interesting difference could possibly be explained by different conformations of the three polymers in solution perhaps induced by the different side chain stereochemistries.

The deprotected polymers, pGlcEMA (**5a**), pGalEMA (**5b**) and pLacEMA (**11**) were obtained following the two different procedures reported in Schemes 3 and 4. In route A, the deprotected monomers **3a**, **3b** and **9** were polymerised with potassium persulfate in a 4:1 mixture of water and methanol at 65 °C for 48 h. In route B, the protected polymers **4a**, **4b** and **10** were deacetylated with sodium methoxide in a 1:1 mixture of chloroform and methanol. In both cases, **5a**, **5b** and **11** were obtained as highly hygroscopic white solids. The three polymers synthesised by polymerisation of the deprotected monomers are denoted by (A); polymers obtained by deacetylating the protected polymers are denoted by (B). pGlcEMA, pGalEMA and pLacEMA were characterised by IR and NMR spectroscopy, optical activity measurements and elemental analysis. All six deprotected polymers synthesised were soluble only in water and the resulting aqueous solutions tended to foam on agitation, indicating surface activity. As will be shown in chapter 4, dynamic light scattering measurements revealed that pGlcEMA and pGalEMA tended to aggregate in aqueous media.

The ^{1}H-NMR spectra of the polymers prepared by each route showed noticeable differences, as reported in Figures 4-6. It can be seen that despite exhaustive deprotecting conditions, signals due to methyl protons of the protecting acetyl groups can still be seen in the range δ 2.0-2.2 ppm in the spectra of polymers obtained by route B, while no trace of these signals is present in the spectra of **5a-A**,

5b-A and **11-A**. As expected, the observed difference was more pronounced for the polymer containing the disaccharide residue. The presence in polymers-B of some residual protecting groups was also shown in the ^{13}C-NMR spectra by peaks in the range δ 22.0-24.0 ppm (see Figure 7 for ^{13}C-NMR spectra of pGalEMA-A and pGalEMA-B). Moreover, microanalyses of these polymers typically gave a higher carbon percentage than expected (Table 2). In only one case was the result in agreement with the calculated composition. It should be noted that deprotection could not be continued to completion due to precipitation of the partially deprotected polymers from solution. Moreover, cleavage of the methacrylate ester bonds could occur during the deacetylation of the pAcGlcEMA, pAcGalEMA and pAcLacEMA, leading to polymers containing methyl ester and/or acid functionalities in the side chains. While there is no clear evidence to support this, the ^{1}H-NMR spectra of polymers prepared by route B generally contain many more weak signals just above the baseline indicating more impurities than in the corresponding compounds synthesised by polymerisation of the deprotected monomers (Figures 4-6). Some of these resonances are in regions where the –OMe signals of methyl methacrylate units would be expected (δ ≈ 3.5 ppm).

Absolute molecular weights of the deprotected polymers were determined by SEC in aqueous solution and the obtained values are reported in Table 3. The refractive index increments, dn/dc, were determined by means of a differential refractometer, working in the same experimental conditions used for the SEC measurements. Values of 0.131 ml g^{-1}, 0.140 ml g^{-1}, 0.150 ml g^{-1}, 0.126 ml g^{-1}, 0.146 ml g^{-1} and 0.142 ml g^{-1} for pGlcEMA-A, pGalEMA-A, pLacEMA-A, pGlcEMA-B, pGalEMA-B and pLacEMA-B, respectively, were obtained.

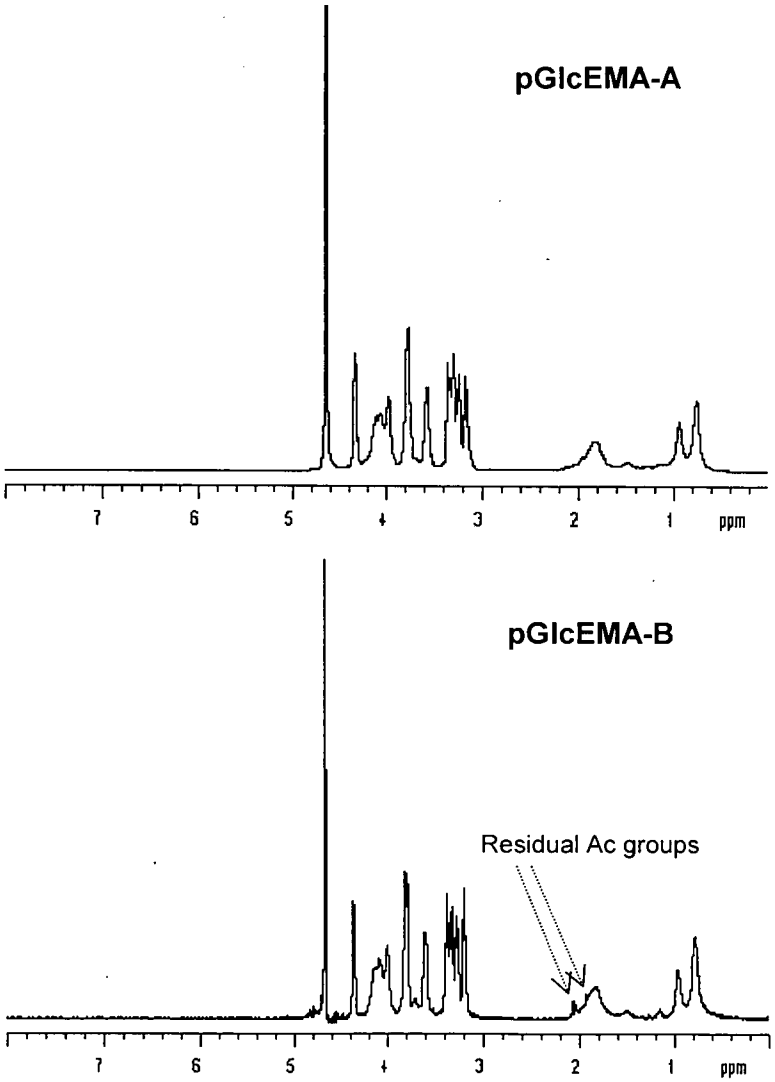


Figure 4 ^{1}H-NMR spectra of pGlcEMA-A and pGlcEMA-B.

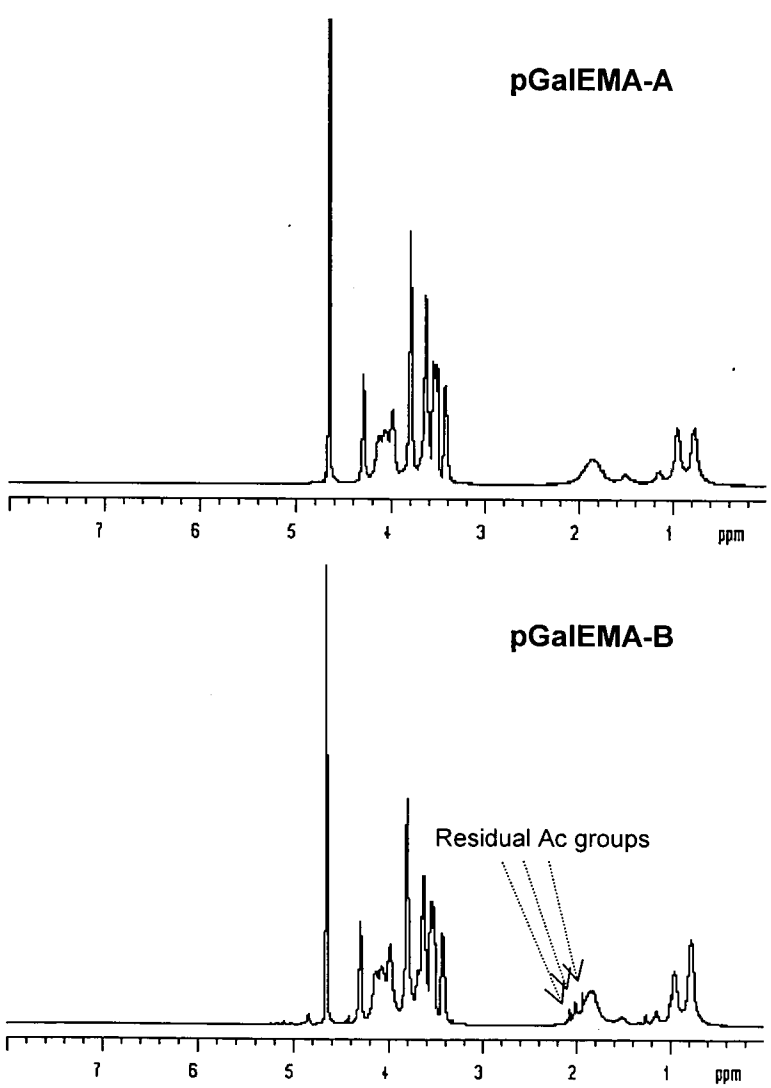


Figure 5 ^{1}H-NMR spectra of pGalEMA-A and pGalEMA-B.

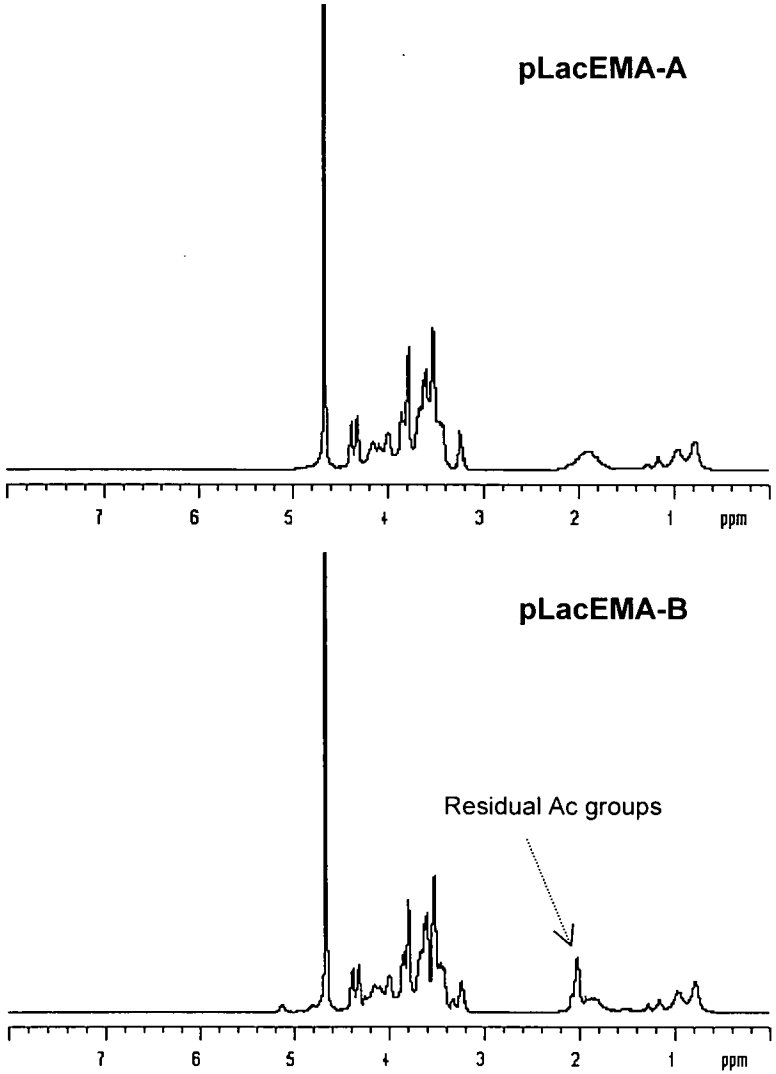


Figure 6 ^{1}H-NMR spectra of pLacEMA-A and pLacEMA-B.

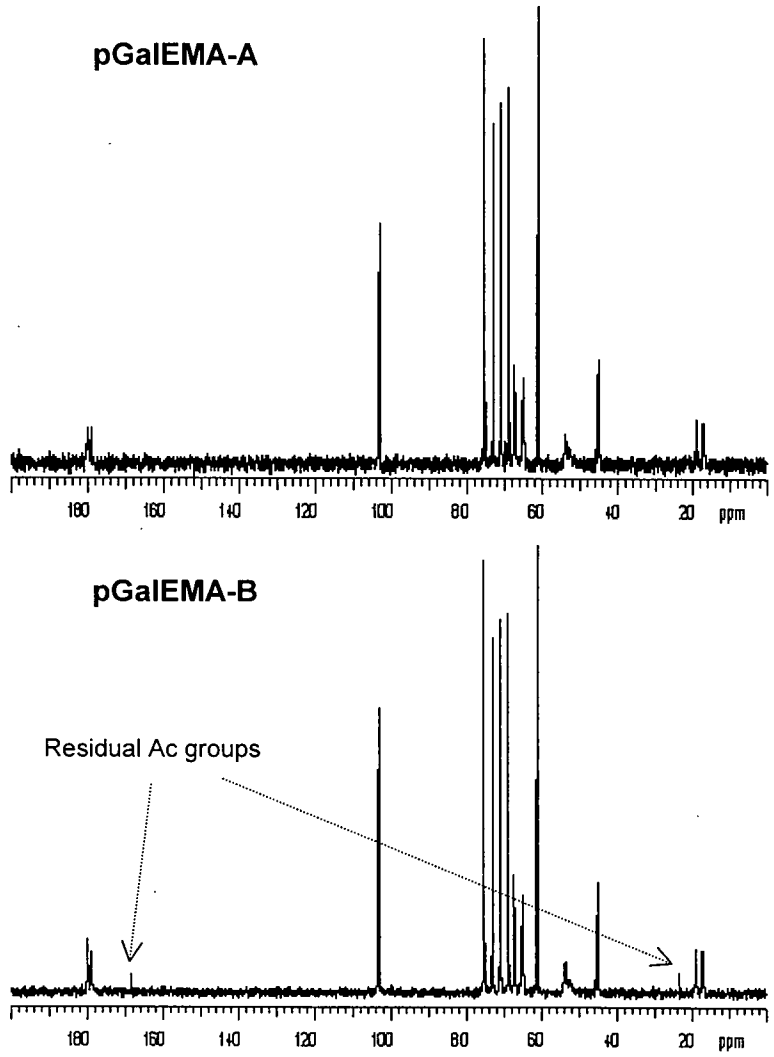


Figure 7 ^{13}C-NMR spectra of pGalEMA-A and pGalEMA-B.

The average molar mass values of **5b-A** are much higher than those of **5a-A**, which may be indicative of a tendency of the former to aggregate more strongly in aqueous solution (it was noticed that aqueous solutions of the deprotected polymers had a tendency to froth). However, more surprising is the difference in average molar mass values of deprotected polymers prepared by route B compared to their parent polymers (compare **5a-B** with **4a**, **5b-B** with **4b** and **11-B** with **10**). While a reduction in molecular weight is expected due to removal of four acetyl groups per sugar residue, this alone cannot account for the differences observed in M_n and M_w between **5-B** and **4** or **11-B** and **10**. Scission of polymer chain C-C bonds during deprotection is an unlikely explanation. The average molar mass values for **5-B** and **11-B** polymers may in fact be artefacts of the polymer deprotection procedure. During the deprotection, a white polymeric precipitate is produced, which is collected as the product. Since the chains become less soluble in the reaction medium as the extent of deprotection of each chain increases, precipitation of shorter chains is likely to occur before longer chains. Therefore, we may have inadvertently fractionated our polymer samples during deprotection, leaving the longer chains in solution. This serves to highlight the differences between the glycopolymer preparation routes and reinforces our belief in the unreliability of the polymer deprotection route B. It is worth noting that "route B" is the generally adopted synthetic procedure: most of the glycopolymers reported in the literature have been prepared by deprotection of the corresponding polymers.

2.3 EXPERIMENTAL

2.3.1 GENERAL

2,3,4,6-Tetra-*O*-acetyl-α-D-glucopyranosyl bromide (95+%) and 2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranosyl bromide (~95%) were purchased from Sigma and Fluka, respectively. Lactose was purchased from Sigma (95+%). 2-Hydroxyethyl methacrylate (HEMA, 99+%), silver trifluoromethanesulfonate (99+%), acetic anhydride (99+%), potassium persulfate (99+%), pyridine (99.8%), methanol (99.8%) and cation-exchange resin DOWEX 50W×2-200 were purchased from Aldrich. 2,2'-Azoisobutyronitrile (AIBN, 97%) was obtained from BDH Laboratory Supplies. HEMA was purified by distillation under vacuum; dichloromethane (DCM) was distilled from calcium hydride under N_2; all other chemicals were used without further purification.

NMR-spectra were recorded with a Varian Inova 500 spectrometer, operating at 499.78 (^{1}H) and at 125.67 MHz (^{13}C); results of HETCOR and COSY correlation studies have been used in order to assign the observed signals to the hydrogen and carbon atoms of the compound. Infrared spectra were obtained with a Perkin Elmer 1600 Series FTIR spectrometer. Mass spectra were obtained with a Micromass Platform spectrometer and with a Micromass LCT spectrometer, ionisation modes ES+ or ES-. Size exclusion chromatography (SEC) of pAcGlcEMA, pAcGalEMA and pAcLacEMA was performed with a Viscotel 200 + light scattering, in THF, using a Plgel 10 μ MIXED-B column and working with a flow rate of 1.000 ml min^{-1} and an injection volume of 100 μl. Absolute molecular weights of pGlcEMA-A,B, pGalEMA-A,B and pLacEMA-A,B were determined by aqueous size exclusion chromatography coupled to a Water 410 RI detector and a Wyatt DAWN DSP

MALLS, using a TSK GMPW column (30 cm × 7.8 mm id). The mobile phase consisted of 80 % HPLC grade water, 20 % methanol, 0.05 M $NaNO_3$ and 2.5 ml l^{-1} 1.0 M NaOH at a flow rate of 0.8 ml min^{-1}. The MALLS detectors were normalized using 50K pullulan purchased from Gearing Scientific. All polymer samples were prepared as 0.2 – 0.6 wt% solutions and filtered through 0.2 μm filters. Molecular weight averages and polydispersity were calculated using Astra 32 software. The dn/dc values used were previously determined using a model BP-2000-V Brice-Phoenix Visual Laboratory Type Differential Refractometer, Phoenix Precision Instrument Company, Pennsylvania (USA); four solutions of concentration ranging between *ca.* 4% and 15% in 20% methanol, 80% 0.05 M $NaNO_3$ and 2.5 ml l^{-1} 1.0 M NaOH solution were prepared for each polymer and the measurements carried out at 25 °C. Optical Rotations were estimated using a P-1020 Series Polarimeter, JASCO (UK) Ltd. Elemental Analyses were obtained with an Exeter Analytical Inc. CE-440 Elemental Analyzer. Products were freeze-dried using a Christ ALPHA-1-4 freeze-dryer with controller LDC-1M.

2.3.2 X-RAY CRYSTALLOGRAPHY

The diffraction experiment (nominally covering full sphere of the reciprocal space) was carried out at T = 100 K on a SMART 3-circle diffractometer with a 1K CCD area detector,[†] using graphite monochromated Mo-K_α radiation ($\overline{\lambda}$ = 0.71073 Å) and a Cryostream (Oxford Cryosystems) open-flow N_2 gas cryostat.

Crystal data. $C_{20}H_{28}O_{12}$, f.w. 460.42, orthorombic, space group $P2_12_12_1$ (No. 19), a = 8.442(2), b = 12.466(4), c = 21.904(7) Å, V = 2305(1) $Å^3$, Z = 4, D_x = 1.33 g

[†] *SMART and SAINT, Area detector control and integration software. Version 6.0.* Bruker AXS, Madison, Wisconsin, U.S.A., 1999.

ml^{-1}, $\mu = 0.11\ mm^{-1}$. 28 228 reflections were measured ($2\theta < 58°$), of these 3451 were independent and 2631 were Friedel equivalents thereof [5679 and 2430 respectively with $I \geq 2\sigma(I)$]. The structure was solved by direct methods and refined by full-matrix least squares against F^2 of all data, using SHELXTL software.[‡] All non-hydrogen atoms were refined in anisotropic approximation, the H atoms at C(12) isotropically, methyl groups as rigid bodies (with a common refined U for three H atoms); other H atoms were treated in 'riding' model. Final R = 0.037 for data with $I \geq 2\sigma(I)$, w$R(F^2)$ = 0.097 for all data. The absolute configuration could not be determined and was assigned from the known one of the starting material. Full data (excluding structure factors) have been deposited at the Cambridge Crystallographic Data Centre. CCDC reference no. 207/171437.

2.3.3 MONOMER AND POLYMER SYNTHESIS

2-(2',3',4',6'-Tetra-*O*-acetyl-*β*-D-glucosyloxy)ethyl methacrylate, AcGlcEMA (1a)

To a stirred solution of 2,3,4,6-tetra-*O*-acetyl-*α*-D-glucopyranosyl bromide (10 g, 24.3 mmol) and an excess of HEMA (9.5 g, 72.8 mmol) in dry DCM at – 40 °C under N_2 containing 3 Å powdered molecular sieves (12 g), an excess of silver trifluoromethanesulfonate (7.5 g, 29.2 mmol) was added and the reaction mixture stirred in the same conditions for 48 h. After this time the mixture was gradually allowed to reach room temperature and then filtered through celite. After evaporation of the solvent, the residue (10.9 g) was dissolved in a mixture of pyridine (250 ml) and acetic anhydride (100 ml) and stirred under N_2 overnight in order to acetylate

[‡] *SHELXTL, an integrated system for solving, refining and displaying crystal structures from diffraction data, Version 5.10.* Bruker AXS, Madison, Wisconsin, USA, 1997.

unreacted HEMA and so aid purification. After evaporation of the solvent the residue was dissolved in DCM (250 ml), washed with a solution of hydrochloric acid (0.25 M, aliquots of 100 ml), a saturated solution of $NaHCO_3$ (aliquots of 100 ml), brine (100 ml) and dried over Na_2SO_4. The solvent was removed and the residue was then purified by flash chromatography (ethylacetate:hexane, 3:7) to afford the product **1a** (8.3 g, 72%) as colourless plate-like crystals suitable for X-ray analysis; mp 77-79 °C (*lit.*[36] 77.5 °C); $[\alpha]_D^{28}$ -12.8 (*c* 0.14 in $CHCl_3$) (*lit.*[36] $[\alpha]_D^{15}$ -12.7 , 0.015 mol dm^{-3} in $CHCl_3$); (Found: C, 52.11; H, 6.19. $C_{20}H_{28}O_{12}$ requires C, 52.17; H, 6.13%); IR (KBr disc) ν/cm^{-1} 1758 (C=O of *O*-acetyl groups), 1718 (C=O of HEMA moiety), 1635, 1321, 1299 (C=C); δ_H[§] (500 MHz; $CDCl_3$) 1.99 (3H, s, 3 × H-3), 2.00, 2.01, 2.03, 2.07 (3H × 4, s × 4, Ac × 4), 3.68 (1H, ddd, $J_{4',5'}$ 10.5 Hz, $J_{5',6a'}$ 2.5 Hz, $J_{5',6b'}$ 4.5 Hz, H-5'), 3.81 (1H, ddd, $J_{6a,6b}$ 11.5 Hz, $J_{5a,6a}$ 7.0 Hz, $J_{5b,6a}$ 3.5 Hz, H-6_a), 4.04 (1H, ddd, $J_{6a,6b}$ 11.5 Hz, $J_{5a,6b} = J_{5b,6b}$ 5.0 Hz, H-6_b) 4.12 (1H, dd, $J_{6a',6b'}$ 12.2 Hz, $J_{5',6a'}$ 2.5 Hz, H-6'$_a$), 4.25 (1H, dd, $J_{6a',6b'}$ 12.2 Hz, $J_{5',6b'}$ 4.5 Hz, H-6'$_b$), 4.23-4.32 (2H, m, 2 × H-$5_{a,b}$), 4.56 (1H, d, $J_{1',2'}$ 7.9 Hz, H-1'), 4.99 (1H, dd, $J_{1',2'}$ 7.7 Hz, $J_{2',3'}$ 9.7 Hz, H-2'), 5.07 (1H, t, *J* 9.7 Hz, H-4'), 5.18 (1H, t, *J* 9.5 Hz, H-3'), 5.57-5.59 (1H, m, H-1 *Z* to CH_3-C=C), 6.10-6.12 (1H, m, H-1 *E* to CH_3-C=C), NOESY correlation studies allowed the differentiation between H-1(*Z*) and H-1(*E*) due to the long range coupling of H-1(*Z*) with 3 × H-3; NOESY also allowed the differentiation between 2 × H-5 and 2 × H-6 due to the long range coupling of H-6 with H-1'; δ_C (125.67 MHz, decoupled ^1H 500 MHz; $CDCl_3$) 18.2 (C-3); 20.5, 20.6 ($H_3\underline{C}COO$- × 4), 61.8 (C-6'), 63.3 (C-5), 67.4 (C-6), 68.2 (C-4'), 71.0 (C-2'), 71.8 (C-5'), 72.6 (C-3'), 100.7 (C-1'), 125.8 (C-1), 135.9 (C-2), 167.1 (C-4), 169.2, 169.3, 170.2, 170.6 ($H_3C\underline{C}OO$- × 4); LRMS *m/z* (ES+): found 483.5 $(M+Na)^+$, 100%.

[§] See Scheme 3 for numbering.

2-(2',3',4',6'-Tetra-*O*-acetyl-*β*-D-galactosyloxy)ethyl methacrylate, AcGalEMA (1b)

2,3,4,6-Tetra-*O*-acetyl-*α*-D-galactopyranosyl bromide (10 g, 24.3 mmol) and an excess of HEMA (9.5 g, 72.8 mmol) were allowed to react and the resulting product purified as described for **1a**, to afford product **1b** (8.9 g, 80%) as a viscous colourless oil; $[\alpha]_D^{28}$ -5.5 (*c* 0.25 in $CHCl_3$); (Found: C, 51.94; H, 6.03. $C_{20}H_{28}O_{12}$ requires C, 52.17; H, 6.13%); IR (nujol) ν/cm^{-1} 1756 (C=O of acetate groups), 1724 (C=O of HEMA moiety), 1632, 1320, 1295 (C=C); δ_H (500 MHz; $CDCl_3$) 1.96 (3H, s, 3 × H-3), 1.99, 2.02, 2.03, 2.13 (3H × 4, s × 4, Ac × 4), 3.82 (1H, ddd, $J_{6a,6b}$ 11.5 Hz, $J_{5a,6a}$ 7.5 Hz, $J_{5b,6a}$ 4.0 Hz, H-6_a), 3.90 (1H, td, $J_t = J_{5',6a'} = J_{5',6b'}$ 6.5 Hz, $J_d = J_{4',5'}$ 1.1 Hz, H-5'), 4.04 (1H, ddd, $J_{6a,6b}$ 11.5 Hz, $J_{5a,6b} = J_{5b,6b}$ 5.0 Hz, H-6_b), 4.10 (1H, dd, $J_{6a',6b'}$ 10.3 Hz, $J_{5',6a'}$ 6.7 Hz, H-6'$_a$), 4.14 (1H, dd, $J_{6a',6b'}$ 10.9 Hz, $J_{5',6b'}$ 6.5 Hz, H-6'$_b$), 4.26 (1H, ddd, $J_{5a,5b}$ 12.0 Hz, $J_{5a,6a}$ 7.5 Hz, $J_{5a,6b}$ 4.5 Hz, H-5_a), 4.31 (1H, ddd, $J_{5a,5b}$ 12.5 Hz, $J_{5b,6a} = J_{5b,6b}$ 5.0 Hz, H-5_b), 4.52 (1H, d, $J_{1',2'}$ 8.0 Hz, H-1'), 4.98 (1H, dd, $J_{2',3'}$ 10.5 Hz, $J_{3',4'}$ 3.5 Hz , H-3'), 5.19 (1H, dd, $J_{1',2'}$ 8.0 Hz, $J_{2',3'}$ 10.5 Hz, H-2'), 5.36 (1H, dd, $J_{3',4'}$ 3.5 Hz, $J_{4',5'}$ 1.0 Hz, H-4'), 5.57-5.58 (1H, m, H-1 *Z* to CH_3-C=C), 6.10-6.11 (1H, m, H-1 *E* to CH_3-C=C); δ_C (125.67 MHz, decouple ^1H 500 MHz; $CDCl_3$) 18.2 (C-3), 20.5 20.6 ($H_3\underline{C}COO$- × 4), 61.2 (C-6'), 63.4 (C-5), 66.9 (C-4'), 67.3 (C-6), 68.5 (C-2'), 70.6 (C-3'), 70.8 (C-5'), 101.2 (C-1'), 125.8 (C-1), 136.0 (C-2), 167.0 (C-4), 169.3 170.1 170.2 170.3 ($H_3C\underline{C}OO$- × 4); LRMS *m/z* (ES+): 483.3 $(M+Na)^+$, 100%.

2-(*O*-acetyl)ethyl methacrylate, AcEMA (2)

Colourless oil; δ_H (200 MHz; CDCl3) 1.92-1.93 (3H, m, CH_3-C=C), 2.05 (3H, s, CH_3-C=O), 4.26-4.36 (4H, m, -CH_2-CH_2-), 5.56-5.58 (1H, m, vinyl proton *Z* to

CH_3-C=C), 6.10 (1H, m, vinyl proton *E* to CH_3-C=C); LRMS *m/z* (ES+): found 195 $(M+Na)^+$, 100%, 211 $(M+K)^+$, 10%.

2-(β-D-Glucosyloxy)ethyl methacrylate, GlcEMA (3a)

1a (2 g, 4.3 mmol) was dissolved in 10 ml of dry methanol and the obtained solution was stirred for 15 min at room temperature under N_2. Then, 1 ml of freshly prepared 0.03 M solution of sodium methoxide in methanol was added. The solution was stirred at room temperature and continuously monitored by TLC (acetonitrile:water, 9:1). The reaction was stopped after 40 min when the formation of the product resulting from the cleavage of the ester bond of HEMA moiety was observed (R_f = 0.2). Cation-exchange resin (DOWEX 50W×2-200) was added in order to bring the pH to neutral and to remove Na^+ cations. The solution was stirred for 15 min before filtering the resin. After evaporation of the solvent, the residue was purified by flash chromatography (chloroform:methanol, 8:2) to afford **3a** as a viscous colourless oil that gave a strongly hygroscopic amorphous colourless solid after freeze-drying (1 g, 80%); $[\alpha]_D^{20}$ - 22.5 (*c* 0.20 in CH_3OH); (Found: C, 48.99; H 7.01. $C_{12}H_{20}O_8$ requires C, 49.31; H 6.90%); IR (KBr disc) ν/cm^{-1} 3364 (OH), 1715 (C=O of HEMA moiety), 1635, 1320, 1298 (C=C); δ_H** (500 MHz; CD_3OD) 1.93 (3H, m, 3 × H-3), 3.17 (1H, dd, $J_{1',2'}$ 8.0 Hz, $J_{2',3'}$ 9.5 Hz, H-2'), 3.25-3.35 (3H, m, H-3', H-4', H-5'), 3.65 (1H, dd, $J_{6a'6b'}$ 12.0 Hz, $J_{5',6a'}$ 5.5 Hz, H-6'$_a$), 3.83 (1H, ddd, $J_{6a,6b}$ 11.5 Hz, $J_{5a,6a}$ 6.0 Hz, $J_{5b,6a}$ 3.5 Hz, H-6$_a$), 3.84 (1H, dd, $J_{6a',6b'}$ 11.7 Hz, $J_{5',6b'}$ 1.7 Hz, H-6'$_b$), 4.09 (1H, ddd, $J_{6a,6b}$ 11.5 Hz, $J_{5a,6b}$ 4.0 Hz, $J_{5b,6b}$ 6.0 Hz, H-6$_b$), 4.29 (1H, ddd, $J_{5a,5b}$ 12.0 Hz, $J_{5a,6a}$ 5.5 Hz, $J_{5a,6b}$ 3.5 Hz, H-5$_a$), 4.30 (1H, d, $J_{1',2'}$ 8.0 Hz, H-1'), 4.34 (1H, ddd, $J_{5a,5b}$ 12.0 Hz, $J_{5b,6a}$ 3.5 Hz, $J_{5b,6b}$ 6.0 Hz, H-5$_b$), 5.61-5.62 (1H, m, H-1 *Z* to CH_3-

** See Scheme 3 for numbering.

C=C), 6.11-6.12 (1H, m, H-1 *E* to CH_3-C=C); δ_C (125.67 MHz, decoupled ^1H 500 MHz; CD_3OD) 18.4 (C-3), 62.7 (C-6'), 65.3 (C-5), 68.6 (C-6), 71.6 (C-3' or C-4' or C-5'), 75.0 (C-2'), 78.0 (2C, C-3' or C-4' or C-5'), 104.6 (C-1'), 126.4 (C-1), 137.6 (C-2), 168.9 (C-4); HRMS *m/z* (ES+): found 3 15.1049 [$(M+Na)^+$. $C_{12}H_{20}O_8$ requires *m/z*, 315.1056].

2-(β-D-Galactosyloxy)ethyl methacrylate, GalEMA (3b)

1b (2 g, 4.3 mmol) was treated using an identical procedure to that described in the preparation of **3a**, to afford **3b** as a viscous colourless oil that gave a strongly hygroscopic amorphous colourless solid after freeze-drying (0.95 g, 75%); $[\alpha]_D^{20}$ - 5.5 (*c* 0.20 in CH_3OH); [Found: C, 48.34; H, 6.96. ($C_{12}H_{20}O_8$ + 0.3 mol H_2O) requires C, 48.42; H, 6.97%]; IR (KBr disc) ν/cm^{-1} 3384 (OH), 1712 (C=O of HEMA moiety), 1636, 1320, 1298 (C=C); δ_H[††] (500 MHz; CD_3OD) 1.93 (3H, m, 3 × H-3), 3.45 (1H, dd, $J_{2',3'}$ 9.7 Hz, $J_{3',4'}$ 3.3 Hz, H-3'), 3.50 (1H, td, $J_t = J_{5',6a'} = J_{5',6b'}$ 5.5 Hz, $J_d = J_{4',5'}$ 1.0 Hz, H-5'), 3.52 (1H, dd, $J_{1',2'}$ 7.5 Hz, $J_{2',3'}$ 10.0 Hz, H-2'), 3.71 (1H, dd, $J_{6a',6b'}$ 11.0 Hz, $J_{5',6a'}$ 5.5 Hz, H-6'$_a$), 3.74 (1H, dd, $J_{6a',6b'}$ 11.0 Hz, $J_{5',6b'}$ 7.0 Hz, H-6'$_b$), 3.82 (1H, dd, $J_{3',4'}$ 3.3 Hz, $J_{4',5'}$ 0.7 Hz, H-4'), 3.84 (1H, ddd, $J_{6a,6b}$ 12.0 Hz, $J_{5a,6a}$ 6.0 Hz, $J_{5b,6a}$ 4.0 Hz, H-6$_a$), 4.10 (1H, ddd, $J_{6a,6b}$ 12.0 Hz, $J_{5a,6b}$ 3.7 Hz, $J_{5b,6b}$ 6.0 Hz, H-6$_b$), 4.26 (1H, d, $J_{1',2'}$ 7.5 Hz, H-1'), 4.31 (1H, ddd, $J_{5a,5b}$ 12.0 Hz, $J_{5a,6a}$ 6.0 Hz, $J_{5a,6b}$ 3.5 Hz, H-5$_a$), 4.35 (1H, ddd, $J_{5a,5b}$ 12.0 Hz, $J_{5b,6a}$ 3.5 Hz, $J_{5b,6b}$ 6.0 Hz, H-5$_b$), 5.61-5.62 (1H, m, H-1 *Z* to CH_3-C=C), 6.12 (1H, m, H-1 *E* to CH_3-C=C); δ_C (125.67 MHz, decoupled ^1H 500 MHz; CD_3OD) 18.4 (C-3), 62.5 (C-6'), 65.3 (C-5), 68.5 (C-6), 70.3 (C-4'), 72.4 (C-2'), 74.9 (C-3'), 76.7 (C-5'), 105.3 (C-1'), 126.4 (C-1), 137.7 (C-2), 168.8 (C-4); HRMS *m/z* (ES+): found 315.1065 [$(M+Na)^+$. $C_{12}H_{20}O_8$ requires *m/z*, 315.1056].

[††] See Scheme 3 for numbering.

Poly[2-(2',3',4',6'-tetra-*O*-acetyl-*β*-D-glucosyloxy)ethyl methacrylate], pAcGlcEMA (4a)

A solution of **1a** (2 g, 4.3 mmol) and AIBN (20 mg, 1 wt%) in chloroform was degassed by bubbling N_2 through for 15 min at room temperature. Then, the flask was sealed and the polymerisation carried out at 65 °C for 48 h. The resulting viscous solution was poured in to an excess of diethyl ether (10:1) in order to precipitate the polymer. The product was then purified by re-precipitation from chloroform in diethyl ether to obtain **4a** (1.9 g, 95%) as a white solid; $[\alpha]_D^{22}$ -11.7 (*c* 0.11 in $CHCl_3$) (*lit.*[36] –10.3, 0.015 mol l^{-1} in benzene); [Found: C, 52.03; H, 6.17. $(C_{20}H_{28}O_{12})_m$ requires C, 52.17; H 6.13%]; IR (KBr disc) ν/cm^{-1} 2960 (-CH_2-), 1756 (C=O of *O*-acetyl groups); δ_H (500 MHz; $CDCl_3$) 0.75-1.15 (3H, br m, CH_3-C), 1.60-2.27 (2H, br, -CH_2-), 2.01, 2.03, 2.06, 2.10 (12H, s × 4, Ac × 4), 3.70-3.88, 3.95-4.07, 4.10-4.25, 4.99-5.03, 5.08-5.18, 5.20-5.28 (10H, protons of the carbohydrate residue and the methylene groups of the side chains), 4.66 (1H, br, anomeric proton); δ_C (125.67 MHz, decoupled ^1H 500 MHz; $CDCl_3$) 16.5 18.2 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr) and *meso*-racemic (mr) triads, respectively), 20.6, 20.7, 20.8 (4C, $H_3\underline{C}COO$- × 4), 45.2 (1C, -CH_2-), 52.2-55.6 (1C, br, CH_3-$\underline{C}$-), 61.1 (1C, carbohydrate residue of the side chains), 63.6 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 66.6 (1C, -$\underline{C}H_2$-OCO), 67.3, 69.1, 70.6, 70.8 (4C, carbohydrate residue of the side chain), 100.5 (1C, anomeric carbon), 169.3, 170.0, 170.3 (4C, $H_3C\underline{C}OO$- × 4), 176.5, 176.9-177.8 (1C, br, -$C\underline{C}OO$-, *meso*-racemic (mr) and racemic-racemic (rr) triads, respectively).

Poly[2-(2',3',4',6'-tetra-*O*-acetyl-*β*-D-galactosyloxy)ethyl methacrylate], pAcGalEMA (4b)

1b (2 g, 4.3 mmol) was polymerised in an identical procedure to that described for **4a**, to obtain **4b** (1.86 g, 93%) as a white solid; $[\alpha]_D^{22}$ -12.2 (*c* 0.11 in $CHCl_3$); [Found: C, 51.92; H, 6.26. $(C_{20}H_{28}O_{12})_m$ requires C, 52.17; H 6.13%]; IR (KBr disc) ν/cm^{-1} 2960(-CH_2-), 1752 (C=O of *O*-acetyl groups); δ_H (500 MHz; $CDCl_3$) 0.78-1.16 (3H, br m, CH_3-C), 1.63-2.20 (2H, br, -CH_2-), 1.98, 2.05, 2.06, 2.16 (12H, s × 4, Ac × 4), 3.71-3.83, 3.95-4.09, 4.10-4.24, 5.05-5.21, 5.37-5.43 (10H, protons of the carbohydrate residue and the methylene groups of the side chains), 4.62 (1H, br, anomeric proton); δ_C (125.67 MHz, decoupled ^1H 500 MHz; $CDCl_3$) 17.4, 18.9 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr) and *meso*-racemic (mr) triads, respectively), 20.8, 20.9, 21.0 (4C, $H_3\underline{C}COO$- × 4), 45.2 (1C, -CH_2-), 52.5-56.1 (1C, br, CH_3-$\underline{C}$-), 61.3 (1C, carbohydrate residue of the side chains), 63.6 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 66.6 (1C, -$\underline{C}H_2$-OCO), 67.3, 68.9, 70.9, 71.0 (4C, carbohydrate residue of the side chain), 101.1 (1C, anomeric carbon), 169.5, 170.2, 170.5 (4C, $H_3C\underline{C}OO$- × 4), 176.5, 177.1-177.8 (1C, br, -$C\underline{C}OO$-, *meso*-racemic (mr) and racemic-racemic (rr) triads, respectively).

Poly[2-(*β*-D-glucosyloxy)ethyl methacrylate], pGlcEMA (5a)

Synthesis A

A solution of **3a** (1 g, 3.4 mmol) and $K_2S_2O_8$ (25 mg, 2.5 wt%) in a mixture of high purity water:methanol (4:1) was degassed by bubbling N_2 through for 15 min at room temperature. Then, the flask was sealed and the polymerisation was carried out at 65 °C for 48 h. The resulting solution was freeze-dried and the recovered polymer was then purified by dialysis against water (Dialysis Tubing-Visking, Size 20 Inf.

Dia. 18/32" – 14.3 mm : 30 M, MWCO –12-14,000 Daltons) for 1 week. The solution was freeze-dried to afford **5a** (0.82 g, 82%) as a white hygroscopic solid; $[\alpha]_D^{22}$ -15.2 (*c* 0.12 in H_2O); {Found: C, 48.82; H 6.92. [$(C_{12}H_{20}O_8)_m$ + 0.1 mol H_2O] requires C, 49.01; H 6.92%}; IR (KBr disc) ν/cm^{-1} 3446 (OH), 1718 (C=O of HEMA moiety); δ_H (500 MHz; D_2O) 0.73-1.22 (3H, br m, CH_3-C), 1.78-2.22 (2H, br, -CH_2-), 3.25-3.53, 3.67-3.76, 3.86-398, 4.08-4.32 (10H, protons of the carbohydrate residue and of the methylene groups of the side chains), 4.47 (1H, d, *J* 7.5 Hz, anomeric proton); δ_C (125.67 MHz, decoupled ^1H 500 MHz; D_2O) 17.2, 18.8, 21.2 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr), *meso*-racemic (mr) and *meso-meso* (mm) triads, respectively), 45.0 (1C, -$\underline{C}H_2$-), 51.0-53.7 (1C, br, CH_3-$\underline{C}$-), 61.1 (1C, carbohydrate residue of the side chains), 65.1 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 67.3 (1C, -$\underline{C}H_2$-OCO), 69.9, 73.3, 76.0, 76.1 (4C, carbohydrate residue of the side chain), 102.7 (1C, anomeric carbon), 178.9, 179.7-180.0 (1C, br, -C$\underline{C}$OO-, *meso*-racemic (mr) and racemic-racemic (rr) triads, respectively).

Synthesis B

A solution of **4a** (2 g) in $CHCl_3$:CH_3OH (1:1, 8 ml) was stirred at room temperature under N_2 for 15 min. Then, 1 ml of a freshly prepared 1 M solution of sodium methoxide in methanol was added and the formation of a white precipitate was observed after around 30 s. After stirring at room temperature under N_2 for 1h, the solid was filtered, dissolved in water and a cation-exchange resin (DOWEX 50W×2-200) was added in order remove Na^+ cations. The solution was stirred for 15 min before filtering the resin. After purification by dialysis against water, the solution was freeze-dried to afford **5a** (0.80 g, 63%) as a white hygroscopic solid; [Found: C, 51.26; H 6.69. $(C_{12}H_{20}O_8)_m$ requires C, 49.31; H 6.90%. Assuming that two *O*-acetyl groups were still present after the deprotection: $(C_{16}H_{24}O_{10})_m$ would require C, 51.06;

H 6.43%. Found: as before]; IR (KBr disc) ν/cm^{-1} 3423 (OH), 1718 (C=O of HEMA moiety); δ_H (500 MHz; D_2O) 0.75-1.18 (3H, br m, CH_3-C), 1.72-2.22 (2H, br, -CH_2-), 2.06, 2.17, 2.20 (methyl protons of *O*-acetyl groups still present after the deprotection), 3.31-3.52, 3.72, 3.92, 4.05-4.36 (10H, protons of the carbohydrate residue and of the methylene groups of the side chains), 4.48 (1H, d, *J* 7 Hz, anomeric proton); δ_C (125.67 MHz, decoupled ^{1}H 500 MHz; D_2O) 17.3, 18.9 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr) and *meso*-racemic (mr) triads, respectively), 45.0 (1C, -$\underline{C}H_2$-), 51.0-53.8 (1C, br, CH_3-$\underline{C}$-), 61.1 (1C, carbohydrate residue of the side chains), 65.1 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 67.3 (1C, -$\underline{C}H_2$-OCO), 69.9, 73.3, 76.0, 76.1 (4C, carbohydrate residue of the side chain), 102.7 (1C, anomeric carbon), 178.8, 179.8-180.0 (1C, br, -C$\underline{C}$OO-, *meso*-racemic (mr) and racemic-racemic (rr) triads, respectively).

Poly[2-(β-D-galactosyloxy)ethyl methacrylate], pGalEMA (5b)

Synthesis A

3b (1 g, 3.4 mmol) was polymerised and the resulting polymer purified according to the procedure described in the synthesis of **5a** to afford **5b** (0.86 g, 86%) as a white hygroscopic solid; $[\alpha]_D^{22}$ + 8.0 (*c* 0.10 in H_2O); {Found: C, 48.82; H 6.92. [$(C_{12}H_{20}O_8)_m$ + 0.1 mol H_2O] requires C, 49.01; H 6.92%}; IR (KBr disc) ν/cm^{-1} 3447 (OH), 1718 (C=O of HEMA moiety); δ_H (500 MHz; D_2O) 0.82-1.20 (3H, br m, CH_3-C), 1.78-2.35 (2H, br, -CH_2-), 3.52-3.59, 3.62-3.72, 3.72-3.85, 3.89-4.02 4.06-4.35 (10H, protons of the carbohydrate residue and of the methylene groups of the side chains), 4.43 (1H, d, *J* 6.5 Hz, anomeric proton); δ_C (125.67 MHz, decoupled ^{1}H 500 MHz; D_2O) 17.2, 18.9 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr) and *meso*-racemic (mr) triads, respectively), 45.0 (1C, -$\underline{C}H_2$-), 52.0-53.8 (1C, br, CH_3-$\underline{C}$-), 61.1 (1C,

carbohydrate residue of the side chains), 65.0 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 67.1 (1C, -$\underline{C}H_2$-OCO), 68.8, 70.9, 73.0, 75.3 (4C, carbohydrate residue of the side chain), 103.2 (1C, anomeric carbon), 178.9, 179.8-180.0 (1C, br, -C$\underline{C}$OO-, *meso*-racemic (mr) and racemic-racemic (rr) triads, respectively).

Synthesis B

4b (2 g) was deprotected and the resulting polymer purified as for **4a** to afford **5b** (0.83 g, 65%) as a white hygroscopic solid; [Found: C, 51.20; H 6.75. $(C_{12}H_{20}O_8)_m$ requires C, 49.31; H, 6.90%. Assuming that two *O*-acetyl groups were still present after the deprotection $(C_{16}H_{24}O_{10})_m$ would require C, 51.06; H 6.43. Found as before]; IR (KBr disc) ν/cm^{-1} 3445 (OH), 1718 (C=O of HEMA moiety); δ_H (500 MHz; D_2O) 0.76-1.16 (3H, br m, CH_3-C), 1.76-2.24 (2H, br, -CH_2-), 2.07, 2.14, 2.20 (methyl protons of *O*-acetyl groups still present after the deprotection), 3.56, 3.61-4.00, 4.12-4.32 (10H, protons of the carbohydrate residue and of the methylene groups of the side chains), 4.43 (1H, d, *J* 6.5 Hz, anomeric proton); δ_C (125.67 MHz, decoupled ^1H 500 MHz; D_2O) 17.2, 18.8, 20.3 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr), *meso*-racemic (mr) and *meso-meso* (mm) triads, respectively), 23.8 (methyl carbon of *O*-acetyl groups still present after the deprotection), 45.0 (1C, -$\underline{C}H_2$-), 51.4-53.7 (1C, br, CH_3-$\underline{C}$-), 61.1 (1C, carbohydrate residue of the side chains), 65.1 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 67.1 (1C, -$\underline{C}H_2$-OCO), 68.7, 70.9, 73.0, 75.3 (4C, carbohydrate residue of the side chain), 103.2 (1C, anomeric carbon), 178.9, 179.7-180.0 (1C, br, -C$\underline{C}$OO-, *meso*-racemic (mr) and racemic-racemic (rr) triads, respectively), 182.1 (-$\underline{C}$OOH due to cleavage of the ester bond of the side chains).

2',3',4',6'-Tetra-*O*-acetyl-*β*-D-galactopyranosyl(1'→4)-2,3,6-tri-*O*-acetyl-*β*-D-glucopyranose (6)

A suspension of anhydrous sodium acetate (20 g, 0.24 mol) in acetic anhydride (140 ml, 1.48 mol) was warmed until it was gently refluxing. After having turned off the heat source D-lactose (20 g, 55.5 mmol) was added portion by portion at a rate that maintained the boiling without external heating. The reaction mixture was refluxed for 2 h, then cooled and poured into ice-water (300 ml). The yellow oil was extracted by DCM (400 ml) and the organic layer washed with a saturated solution of $NaHCO_3$ (aliquots of 200 ml) and dried over $MgSO_4$. The solvent was removed under vacuum to afford **6** (87% as β anomer) as a crystalline white solid (37.7 g, 100%); $C_{28}H_{38}O_{19}$; mp 96-98 °C; δ_H[‡‡] (500 MHz; $CDCl_3$) 1.95, 2.02, 2.03, 2.04 (3H × 8, s × 8, Ac × 8), 3.75 (1H, ddd, $J_{4,5}$ 10 Hz, $J_{5,6a}$ 4.5 Hz, $J_{5,6b}$ 2 Hz, H-5), 3.81-3.88 (2H, m, H-4, H-5'), 4.06 (1H, dd, $J_{6a',6b'}$ 11.0 Hz, $J_{5',6a'}$ 7.0 Hz, H-6'$_a$), 4.09-4.14 (2H, m, H-6'$_b$, H-6$_a$), 4.43-4.48 (1H, m, H-6$_b$), 4.46 (1H, d, $J_{1',2'}$ 8.0 Hz, H-1'), 4.93 (1H, dd, $J_{2',3'}$ 10.5 Hz, $J_{3',4'}$ 3.5 Hz, H-3'), 5.03 (1H, dd, $J_{1,2}$ 8.5 Hz, $J_{2,3}$ 9.5 Hz, H-2), 5.09 (1H, dd, $J_{1',2'}$ 8.0 Hz, $J_{2',3'}$ 10.5 Hz, H-2'), 5.23 (1H, t, *J* 9.0 Hz, H-3), 5.34 (1H, d, $J_{3',4'}$ 3.5 Hz, H-4'), 5.66 (1H, d, $J_{1,2}$ 8.0 Hz, H-1); δ_C (125.70 MHz, decoupled ^1H 500 MHz; $CDCl_3$) 20.7, 20.8, 20.9, 21.0, 21.1 ($H_3\underline{C}COO$- × 8), 61.0 (C-6'), 61.9 (C-6), 66.8 (C-4'), 69.2 (C-2'), 70.7 (C-2), 70.9 (C-5'), 71.2 (C-3'), 72.8 (C-3), 73.7 (C-5), 75.6 (C-4), 91.5 (C-1), 100.9 (C-1'), 169.1, 169.3, 169.8, 169.9, 170.3, 170.4, 170.5, 170.6 ($H_3C\underline{C}OO$- × 8); LRMS *m/z* (ES+): found 701.1 $(M+Na)^+$, 100%.

‡‡ See Scheme 4 for numbering. The discussed peaks are the ones for the highly predominant β anomer.

2',3',4',6'-Tetra-*O*-acetyl-*β*-D-galactopyranosyl(1'→4)-2,3,6-tri-*O*-acetyl-*α*-D-glucopyranosyl bromide (7)

6 (20 g, 29.5 mmol) was dissolved in dry DCM 125 ml containing 3 Å powdered molecular sieves (10 g). A ready-made solution of HBr in CH_3COOH (30%, 40 ml) was added and the mixture stirred under N_2 overnight. The mixture was then poured in ice-water (200 ml) and the oil extracted by DCM (100 ml). The organic layer was washed with a saturated solution of $NaHCO_3$ (aliquots of 50 ml) and dried over $MgSO_4$. The solvent was removed and the residue was purified by flash chromatography (ethylacetate:dichloromethane, 1:9) to afford the product **7** (14.9 g, 72%) as a white solid; $C_{26}H_{35}BrO_{17}$; mp 136-138 °C; δ_H[§§] (500 MHz; $CDCl_3$) 1.96, 2.03, 2.05, 2.06, 2.09, 2.13, 2.15 (3H × 7, s × 7, Ac × 7), 3.85 (1H, t, *J* 8.75 Hz, H-4), 3.88 (1H, td, J_t 6.0 Hz, J_d 1.5 Hz, H-5'), 4.05-4.21 (4H, m, H-6'$_a$, H-6'$_b$, H-6$_a$, H-5), 4.47-4.52 (1H, m, H-6$_b$), 4.50 (1H, d, $J_{1',2'}$ 8.0 Hz, H-1'), 4.75 (1H, dd, $J_{1,2}$ 4.0 Hz, $J_{2,3}$ 10.0 Hz, H-2), 4.95 (1H, dd, $J_{2',3'}$ 10.5 Hz, $J_{3',4'}$ 3.5 Hz, H-3'), 5.12 (1H, dd, $J_{1',2'}$ 8.0 Hz, $J_{2',3'}$ 10.5 Hz, H-2'), 5.35 (1H, dd, $J_{3',4'}$ 3.5 Hz, $J_{4',5'}$ 1.2 Hz, H-4'), 5.55 (1H, t, *J* 9.5 Hz, H-3), 6.25 (1H, d, $J_{1,2}$ 4.0 Hz, H-1); δ_C (125.70 MHz, decoupled ^1H 500 MHz; $CDCl_3$) 20.5, 20.6, 20.8 ($H_3\underline{C}COO$- × 7), 60.8 (C-6'), 61.0 (C-6), 66.5 (C-4'), 69.0 (C-2'), 69.5 (C-3), 70.7 (C-2), 70.8 (C-5'), 71.0 (C-3'), 72.9 (C-5), 74.9 (C-4), 86.6 (C-1), 100.8 (C-1'), 168.9, 169.2, 170.0, 170.1, 170.2, 170.2, 170.6 ($H_3C\underline{C}OO$- × 7); LRMS *m/z* (ES+): found 720.7, 722.7 $(M+Na)^+$, 100%, 98.86%, respectively.

[§§] See Scheme 4 for numbering.

2-[2'',3'',4'',6''-Tetra-*O*-acetyl-β-D-galactopyranosyl(1''→4')-2',3',6'-tri-*O*-acetyl-β-D-glucosyloxy]ethyl methacrylate, AcLacEMA (8)

7 (20 g, 28.6 mmol) and an excess of HEMA (11.1 g, 85.3 mmol) were allowed to react as described for **1a**. After removal of the solvent, the residue was purified by flash chromatography (dichloromethane:ethylacetate, 9:1) to afford product **8** (7.9 g, 37%) as a white, amorphous solid; mp 50-52 °C; $[\alpha]_D^{23}$ -16.9 (*c* 0.10 in $CHCl_3$); [Found: C, 50.82; H, 5.96. ($C_{32}H_{44}O_{20}$ + 0.25 mol H_2O) C, 51.03; H, 5.96%]; IR (KBr disc) ν/cm^{-1} 1753 (C=O of *O*-acetyl groups + C=O of HEMA moiety), 1636, 1321, 1298 (C=C); δ_H*** (500 MHz; $CDCl_3$) 1.94 (3H, s, 3 × H-3), 1.96, 2.00, 2.04, 2.04, 2.06, 2.12, 2.15 (3H × 7, s × 7, Ac × 7), 3.59 (1H, ddd, $J_{4',5'}$ 10.0 Hz, $J_{5',6a'}$ 2.0 Hz, $J_{5',6b'}$ 5.0 Hz, H-5'), 3.76-3.81 (2H, m, H-4', H-6$_a$), 3.86 (1H, t, *J* 7.0 Hz, H-5'') 4.05 (1H, ddd, $J_{5a,6b}$ 5.0 Hz, $J_{5b,6b}$ 3.5 Hz, $J_{6a,6b}$ 11.5 Hz H-6$_b$), 4.05-4.14 (3H, m, H-6'$_a$, H-6'$_b$, H-6''$_a$), 4.23-4.31 (2H, m, 2 × H-5$_{a,b}$), 4-46-4.49 (1H, m, H-6''$_b$), 4.47 (1H, d, $J_{1'',2''}$ 8.0 Hz, H-1''), 4.52 (1H, d, $J_{1',2'}$ 8.0 Hz, H-1'), 4.90 (1H, dd, $J_{1',2'}$ 7.5 Hz, $J_{2',3'}$ 9.5 Hz, H-2'), 4.95 (1H, dd, $J_{2'',3''}$ 10.5 Hz, $J_{3'',4''}$ 3.5 Hz, H-3''), 5.10 (1H, dd, $J_{1'',2''}$ 8.0 Hz, $J_{2'',3''}$ 10.5 Hz, H-2''), 5.18 (1H, t, *J* 9.5, H-3'), 5.34 (1H, d, $J_{3'',4''}$ 3.4 Hz, H-4''), 5.58-5.59 (1H, m, H-1 *Z* to CH_3-C=C), 6.11 (1H, s, broad, H-1 *E* to CH_3-C=C); NOESY allowed the differentiation between H-1' and H-1'' due to the long range coupling of H-6 with H-1'; δ_C (125.70 MHz, decoupled ^{1}H 500 MHz; $CDCl_3$) 18.5 (C-3); 20.8, 20.9, 21.0, 21.1 ($H_3\underline{C}COO$- × 7), 61.0 (C-6'), 62.2 (C-6''), 63.6 (C-5), 66.8 (C-4'), 67.7 (C-6), 69.3 (C-2''), 70.9 (C-5''), 71.2 (C-3''), 71.7 (C-2'), 72.9 (C-5'), 73.0 (C-3'), 76.5 (C-4''), 100.8 (C-1'), 102.4 (C-1''), 126.2 (C-1), 136.3 (C-2), 167.4 (C-4), 169.3, 169.8, 170.0, 170.3, 170.4, 170.6, 170.7 ($H_3C\underline{C}OO$- × 7); HRMS *m/z* (ES+): found 771.2324 [$(M+Na)^+$. $C_{12}H_{20}O_8$ requires 771.2324].

*** See Scheme 4 for numbering.

2-[β-D-Galactopyranosyl(1''→4')-β-D-glucosyloxy]ethyl methacrylate, LacEMA (9)

8 (1 g, 1.3 mmol) was treated using an identical procedure to that described in the preparation of **3a**. After evaporation of the solvent, the residue was purified by flash chromatography (dichloromethane:methanol:acetic acid:water, 70:20:1:2) and the solvent removed under vacuum first and by freeze-drying after to afford **9** (0.25 g, 41%) as a strongly hygroscopic amorphous white solid; $[\alpha]_D^{24}$ – 7.6 (*c* 0.10 in H_2O); (Found: C, 47.31; H, 6.46. $C_{18}H_{30}O_{13}$ requires C, 47.58; H, 6.65%); IR (KBr disc) ν/cm^{-1} 3 376 (OH), 1 718 (C=O of HEMA moiety), 1 636, 1 320, 1 299 (C=C); δ_H[†††] (500 MHz; CD_3OD) 1.92 (3H, m, 3 × H-3), 3.31 (1H, t, $J_t = J_{1',2'} = J_{2',3'}$ 8.2 Hz, H-2'), 3.52 (1H, dd, $J_{2',3'}$ 8.0 Hz, $J_{3',4'}$ 10.0 Hz, H-3'), 3.63-3.66 (2H, m, H-5'', H-6''$_a$), 3.67-3.72 (2H, m, H-4', H-6''$_b$), 3.73-3.81 (2H, m, H-5', H-6'$_a$), 3.76 (1H, t, $J_t = J_{1'',2''} = J_{2'',3''}$ 8.2 Hz, H-2''), 3.79 (1H, dd, $J_{2'',3''}$ 11.5 Hz, $J_{3'',4''}$ 3.5 Hz, H-3''), 3.91 (1H, d, $J_{3'',4''}$ 3.50Hz, H-4''), 3.94 (1H, dd, $J_{5',6a'}$ 2.0 Hz, $J_{6a',6b'}$ 12.5 Hz, H-6'$_b$), 3.98 (1H, ddd, $J_{6a,6b}$ 11.5 Hz, $J_{5a,6a}$ 7.5 Hz, $J_{5b,6a}$ 3.0 Hz, H-6$_a$), 4.15 (1H, ddd, $J_{6a,6b}$ 11.5 Hz, $J_{5a,6b}$ 8.5 Hz, $J_{5b,6b}$ 2.5 Hz, H-6$_b$), 4.35-4.39 (2H, m, 2 × H-5),), 4.43 (1H, d, $J_{1'',2''}$ 8.0 Hz, H-1''), 4.53 (1H, d, $J_{1',2'}$ 8.0 Hz, H-1'), 5.72 (1H, m, H-1 *Z* to CH_3-C=C), 6.16 (1H, m, H-1 *E* to CH_3-C=C); δ_C (125.67 MHz, decoupled ^1H 500 MHz; CD_3OD) 17.5 (C-3), 60.1 (C-4''), 64.3 (C-5), 66.1 (C-6'), 68.2 (C-5''), 68.7 (C-6), 70.6 (C-3'), 71.8 (C-3''), 72.7 (C-5'), 72.8 (C-2''), 74.4 (C-4'), 75.5 (C-6''), 78.3 (C-2'), 102.5 (C-1'), 103.0 (C-1''), 127.2 (C-1), 135.9 (C-2), 169.8 (C-4); HRMS *m/z* (ES+): found 477.1589 [$(M+Na)^+$. $C_{18}H_{30}O_{13}$ requires *m/z*, 477.1584].

[†††] See Scheme 4 for numbering.

Poly{2-[2'',3'',4'',6''-tetra-*O*-acetyl-*β*-D-galactopyranosyl(1''→4')-2',3',6'-tri-*O*-acetyl-*β*-D-glucosyloxy]ethyl methacrylate}, pAcLacEMA (10)

8 (1 g, 1.3 mmol) was polymerised in an identical procedure to that described for **4a**, to obtain **10** (0.72 g, 72%) as a white solid; $[\alpha]_D^{23}$ –31.0 (*c* 0.12 in $CHCl_3$); {Found: C, 50.71; H, 5.90. [$(C_{32}H_{44}O_{20})_m$ + 0.5 mol H_2O] requires C, 50.73; H 5.99%}; IR (KBr disc) ν/cm^{-1} 2960 (-CH_2-), 1753 (C=O of acetate groups); δ_H (500 MHz; $CDCl_3$) 0.75-1.10 (3H, br m, CH_3-C), 1.70-2.25 (2H, br, -CH_2-), 1.94, 1.96, 1.99, 2.00, 2.05, 2.12, 2.15 (21H, s × 7, Ac × 7), 3.70, 3.88-3.96, 4.09-4.12, 4.48-4.62, 4.88, 5.07, 5.18, 5.35 (16H, protons of the carbohydrate residue and the methylene groups of the side chains), 4.52 (1H, br, anomeric proton), 5.02 (1H, broad, anomeric proton); δ_C (125.67 MHz, decoupled ^{1}H 500 MHz; $CDCl_3$) 16.5 18.2 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr) and *meso*-racemic (mr) triads, respectively), 20.5, 20.6, 20.8 (7C, $H_3\underline{C}COO$- × 7), 45.0 (1C, -CH_2-), 52.5-55.5 (1C, br, CH_3-$\underline{C}$-), 60.8 (1C, carbohydrate residue of the side chains), 62.5 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 63.2 (1C, -$\underline{C}H_2$-OCO), 66.8, 67.3, 69.4, 70.7, 71.2, 71.7, 73.0, 73.2 (9C, carbohydrate residue of the side chain), 100.3, 101.0 (2C, anomeric carbons), 169.2, 169.5, 169.8, 170.0, 170.2 (7C, $H_3C\underline{C}OO$- × 7), 176.5-178.2 (1C, br, -$C\underline{C}OO$-).

Poly{2-[*β*-D-galactopyranosyl(1''→4')-*β*-D-glucosyloxy]ethyl methacrylate}, pLacEMA (11)

Synthesis A

9 (1 g, 3.4 mmol) was polymerised and the resulting polymer purified according to the procedure described in the synthesis of **5a** to afford **11** (0.36 g, 59%) as a white hygroscopic solid; $[\alpha]_D^{24}$ – 12.7 (*c* 0.11 in H_2O); {Found: C, 46.79; H 6.82.

[$(C_{18}H_{30}O_{13})_m$ + 0.3 mol H_2O] requires C, 47.02; H 6.71%}; IR (KBr disc) ν/cm^{-1} 3417 (OH), 1719 (C=O of HEMA moiety); δ_H (500 MHz; D_2O) 0.72-1.16 (3H, br m, CH_3-C), 1.78-2.25 (2H, br, -CH_2-), 3.22, 3.36, 3.44-3.57, 3.62, 3.64-3.69, 3.82-3.85, 3.99-4.22 (16H, protons of the carbohydrate residue and of the methylene groups of the side chains), 4.33 (1H, br, anomeric proton), 4.40 (1H, br, anomeric proton); δ_C (125.67 MHz, decoupled ^1H 500 MHz; D_2O) 16.8, 17.9 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr) and *meso*-racemic (mr) triads, respectively), 45.0 (1C, -$\underline{C}H_2$-), 53.0-54.1 (1C, br, CH_3-$\underline{C}$-), 60.2, 61.3 (2C, carbohydrate residue of the side chains), 65.2 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 67.1 (1C, -$\underline{C}H_2$-OCO), 68.8, 70.9, 72.9, 73.0, 74.5, 75.3, 75.6, 78.8 (10C, carbohydrate residue of the side chain), 102.2 (1C, anomeric carbon), 103.2 (1C, anomeric carbon), 175.1, 177.8-180.0 (1C, br, -C$\underline{C}$OO-, *meso*-racemic (mr) and racemic-racemic (rr) triads, respectively).

Synthesis B

10 (0.40 g) was deprotected and the resulting polymer purified as for **4a** to afford **11** (0.18 g, 76%) as a white hygroscopic solid; $[\alpha]_D^{24}$ – 6.5 (*c* 0.10 in H_2O); [Found: C, 54.31; H 6.38. $(C_{18}H_{30}O_{13})_m$ requires C, 47.58; H, 6.65%]; IR (KBr disc) ν/cm^{-1} 3423 (OH), 1752 (C=O of residual *O*-acetyl groups), 1718 (C=O of HEMA moiety); δ_H (500 MHz; D_2O) 0.75-1.14 (3H, br m, CH_3-C), 1.73-2.18 (2H, br, -CH_2-), 1.96, 2.02, 2.08 (methyl protons of *O*-acetyl groups still present after the deprotection), 3.24, 3.33, 3.43-3.52, 3.60-3.67, 3.79-3.86, 3.99-4.5 (16H, protons of the carbohydrate residue and of the methylene groups of the side chains), 4.31 (1H, br, anomeric proton), 4.38 (1H, br, anomeric proton); δ_C (125.67 MHz, decoupled ^1H 500 MHz; D_2O) 16.7, 18.0 (1C, br, $\underline{C}H_3$-C, racemic-racemic (rr) and *meso*-racemic (mr) triads, respectively), 20.7 (methyl carbon of *O*-acetyl groups still present after the deprotection), 44.9 (1C, -$\underline{C}H_2$-), 53.4-54.4 (1C, br, CH_3-$\underline{C}$-), 60.3, 61.1 (2C,

carbohydrate residue of the side chains), 65.0 (1C, -$\underline{C}H_2$-O-carbohydrate residue), 67.3 (1C, -$\underline{C}H_2$-OCO), 68.7, 71.1, 72.7, 72.9, 74.6, 75.0, 75.5, 78.7 (10C, carbohydrate residue of the side chain), 102.5 (1C, anomeric carbon), 103.1 (1C, anomeric carbon), 174.5-179.9 (1C, br, -C$\underline{C}$OO-).

2.4 CONCLUSIONS

Monomeric and polymeric methacrylate derivatives bearing β-D-glucopyranoside, β-D-galactopyranoside and β-D-lactoside residues have been successfully synthesised in an efficient stereo-controlled manner and, for the compounds containing the monosaccharide units, in high yields. All products have been fully characterised. Critically, it has been shown that deprotection of the polymers according to the methods used for all previous syntheses of poly(glycoacrylates) and –methacrylates[36,38-40] results in an incomplete deacetylation and yields products of ill-defined composition. Instead, fully deacetylated, well-defined and pure materials can be obtained by the *novel* method of polymerisation of the deprotected monomers. Absolute number- and weight-average molecular weights have been determined for all the polymers synthesised.

The thorough comparison of all functional properties of these glycopolymers, including binding of the deprotected glycopolymers to specific receptor proteins (lectins) to assess their feasibility for biomedical applications, are reported in the next chapters.

2.5 BIBLIOGRAPHICAL REFERENCES

[1] T. Miyata, K. Nakamae, *Trends Polym. Sci.* **1997**, *5*, 198-206.

[2] V. Horejsi, P. Smolek, J. Kocourek, *Biochim. Biophys. Acta* **1978**, *538*, 293-298.

[3] K. Kobayashi, N. Kakishita, M. Okada, T. Akaike, S. Kitazawa, *Makromol. Chem., Rapid Commun.* **1993**, *14*, 293-302.

[4] T. Nakaya, M. Memita, M. Kang, *Macromol. Reports* **1993**, *A30(suppl. 5)*, 349-355.

[5] G. Wulff, G. Clarkson, *Macromol. Chem. Phys.* **1994**, *195*, 2603-2610.

[6] T. Furuike, N. Nishi, S. Tokura, S.-I. Nishimura, *Chem. Lett.* **1995**, 823-824.

[7] K. Matsuoka, S.-I. Nishimura, *Macromolecules* **1995**, *28*, 2961-2968.

[8] G. Wulff, J. Schmid, T. Venhoff, *Macromol. Chem. Phys.* **1996**, *197*, 259-274.

[9] K. Kobayashi, A. Tsuchida, T. Usui, T. Akaike, *Macromolecules* **1997**, *30*, 2016-2020.

[10] K. Ohno, Y. Tsujii, T. Miyamoto, T. Fukuda, M. Goto, K. Kobayashi, T. Akaike, *Macromolecules* **1998**, *31*, 1064-1069.

[11] K. Ohno, Y. Tsujii, T. Fukuda, *J. Polym. Sci. Part A: Polym. Chem.* **1998**, *36*, 2473-2481.

[12] K. Ohno, Y. Izu, S. Yamamoto, T. Miyamoto, T. Fukuda, *Macromol. Chem. Phys.* **1999**, *200*, 1619-1625.

[13] G. Wulff, H. Schmid, L. Zhu, *Macromol. Chem. Phys.* **1999**, *200*, 774-782.

[14] X.-C. Liu, J. S. Dordick, *J. Polym. Sci. Part A: Polym. Chem.* **1999**, *37*, 1665-1671.

[15] K. Hashimoto, T. Ohsawa, H. Saito, *J. Polym. Sci. Part A: Polym. Chem.* **1999**, *37*, 2773-2779.

[16] T. Nukada, A. Berces, M. Z. Zgierski, D. M. Whitfield, *J. Am. Chem. Soc.* **1998**, *120*, 13291-13295.

[17] A. Takasu, T. Niwa, H. Itou, Y. Inai, T. Hirabayashi, *Macromol. Rapid Commun.* **2000**, *21*, 764-769.

[18] G. B. Sigal, M. Mammen, G. Dahmann, G. M. Whitesides, *J. Am. Chem. Soc.* **1996**, *118*, 3789-3800.

[19] B. G. Davis, *J. Chem. Soc., Perkin Trans. 1* **1999**, 3125-3237.

[20] M. Okada, *Progr. Polym. Sci.* **2001**, *26*, 67-104.

[21] J. Kopeček, P. Kopečkova, H. Brøndsted, R. Rathi, B. Ríhová, P.-Y. Yeh, K. Ikesue, *J. Control. Release* **1992**, *19*, 121-130.

[22] P. Kopečkova, R. Rathi, S. Takada, B. Ríhová, M. M. Berenson, J. Kopeček, *J. Control. Release* **1994**, *28*, 211-222.

[23] K. Kobayashi, A. Kobayashi, S. Tobe, T. Akaike, in *Neoglycoconjugates: Preparation and Applications* (Eds.: Y. C. Lee, R. T. Lee), Academic Press, London, **1994**, pp. 262-284.

[24] A. Tsuchida, K. Kobayashi, N. Matsubara, T. Muramatsu, T. Suzuki, Y. Suzuki, *Glycoconjugate J.* **1998**, *15*, 1047-1054.

[25] K. Ohno, T. Fukuda, H. Kitano, *Macromol. Chem. Phys.* **1998**, *199*, 2193-2197.

[26] R. Bahulekar, T. Tokiwa, J. Kano, T. Matsumura, I. Kojima, M. Kodama, *Carbohydr. Polym.* **1998**, *37*, 71-78.

[27] T. Furuike, S. Aiba, T. Suzuki, T. Takahashi, Y. Suzuki, K. Yamada, S.-I. Nishimura, *J. Chem. Soc., Perkin Trans. 1* **2000**, 3000-3005.

[28] S.-H. Kim, M. Goto, C.-S. Cho, T. Akaike, *Biotech. Lett.* **2000**, *22*, 1049-1057.

[29] E. J. Gordon, J. E. Gestwicki, L. E. Strong, L. L. Kiessling, *Chem. Biol.* **2000**, *7*, 9-16.

[30] M.-G. Baek, R. Roy, *Biomacromolecules* **2000**, *1*, 768-770.

[31] N. Nagahori, S.-I. Nishimura, *Biomacromolecules* **2001**, *2*, 22-24.

[32] S. Matsumura, H. Kubokawa, K. Toshima, *Macromol. Chem., Rapid Commun.* **1993**, *14*, 55-58.

[33] T. Mori, S. Fujita, Y. Okahata, *Carbohydr. Res.* **1997**, *298*, 65-73.

[34] M. Ambrosi, A. Batsanov, N. R. Cameron, B. G. Davis, J. A. K. Howard, R. Hunter, *J. Chem. Soc., Perkin Trans. 1* **2002**, 45-52.

[35] S. Kitazawa, M. Okumura, K. Kinomura, T. Sakakibara, *Chem. Lett.* **1990**, 1733-1736.

[36] T. Nakaya, K. Nishio, M. Memita, M. Imoto, *Makromol. Chem., Rapid Commun.* **1993**, *14*, 77-83.

[37] B. Helferich, K. Weis, *Chem. Ber.* **1956**, *89*, 314-321.

[38] Y.-Z. Liang, Z.-C. Li, G.-Q. Chen, F.-M. Li, *Polym. Int.* **1999**, *48*, 739-742.

[39] Y.-Z. Liang, Z.-C. Li, F.-M. Li, *J. Colloid Interface Sci.* **2000**, *224*, 84-90.

[40] Z.-C. Li, Y.-Z. Liang, G.-Q. Chen, F.-M. Li, *Macromol. Rapid Commun.* **2000**, *21*, 375-380.

[41] R. U. Lemieux, A. R. Morgan, *Can. J. Chem.* **1965**, *43*, 2190-2198.

[42] R. T. Lee, Y. C. Lee, *Carbohydr. Res.* **1995**, *271*, 131-136.

[43] M.-Z. Liu, H.-N. Fan, Z.-W. Guo, Y.-Z. Hui, *Carbohydr. Res.* **1996**, *290*, 233-237.

[44] R. U. Lemieux, *Chem. Can.* **1964**, *16*, 14-18.

[45] G. Wulff, G. Röhle, *Angew. Chem., Int. Ed. Engl.* **1974**, *13*, 157-170.

[46] H. Kondo, S. Aoki, Y. Ichikawa, R. L. Halcomb, H. Ritzen, C. H. Wong, *J. Org. Chem.* **1994**, *59*, 864-877.

[47] N. K. Kochetkov, V. M. Zhulin, E. M. Klimov, N. N. Malysheva, Z. G. Makarova, A. Y. Ott, *Carbohydr. Res.* **1987**, *164*, 241-254.

[48] N. S. Banait, W. P. Jencks, *J. Am. Chem. Soc.* **1991**, *113*, 7951-7958.

[49] C. Chiappe, G. Lo Moro, P. Munforte, *Tetrahedron* **1997**, *53*, 10471-10478.

3

SURFACE ACTIVITY OF PROTECTED GLYCOPOLYMERS

3.1 INTRODUCTION: NUCLEATION BENEATH LANGMUIR FILMS

Langmuir monolayers have attracted and continue to attract wide interest as model systems for biological membranes and for the understanding of the early stages of molecular self-assembly[1].

Recently, the research on Langmuir-Blodgett (LB) films has rapidly increased due to their potential applications in the field of non-linear optics, electronic devices, sensors and so on. In particular, poly(methyl methacrylate) (PMMA) LB films have been significantly studied because of their use in nanolithography and non-linear optics[2,3]. Transfer characteristics onto silicon substrates of several polymethacrylates' films, including poly(ethyl methacrylate) (PEMA) and poly[(2-hydroxypropyl)methacrylate] (PHPMA), have been investigated. All the monolayers showed qualitatively similar transfer properties, gradually passing from an initial Z-type transfer to a Y-type transfer with increasing the transfer surface pressure[4].

On the other hand, the oriented growth of crystalline solids in biological systems, called biomineralisation[5], has inspired numerous studies of crystallisation induced by surfactant films. In Nature, in fact, organic systems, such as aggregates of proteins or phospholipids, are found to promote the growth of one particular polymorph or morphology related to a particular biological function. Langmuir monolayers have been thus shown to act efficiently as matrixes beneath which

heterogeneous, nucleation and growth preferentially occur[6]. This nucleation is often highly specific, opening important applicative possibilities in the fields of crystal morphological engineering[7] and biomineralisation[8].

In order for a film to work effectively as a preferential nucleation site, the interfacial tension between itself and the nucleating species must be low. Moreover, the crystallising material must be adsorbed significantly onto the monolayer and the resultant arrangement of the adsorbed molecules must be able to induce nucleation of the bulk crystal[9]. If these conditions are satisfied, epitaxial growth of the crystalline material beneath the film can be achieved, with transfer of structural information from the 2D monolayer to the 3D crystal. Certainly, if the surfactant is chosen so that the film headgroups mimic perfectly or at least closely the molecular array of a particular plane of the crystallising species, the monolayer will act as a highly efficient and specific nucleator, leading to nucleation bounded by this plane[10]. This template-directed mechanism, based on the exact or close match between the lattice structures of the film and the to-be-grown crystal, has been extensively investigated and employed[11-14]. Landau *et al.*[11,12] were the first to demonstrate that Langmuir films can promote the oriented nucleation of both organic and inorganic crystalline solids. For example, glycine crystals were grown with their {010} faces attached to a monolayer constituted by resolved amino acid surfactants. The ability of the film to induce nucleation was directly related to the degree of lattice match between the {010} glycine faces and the film structure. However, epitaxial growth beneath monolayers has been shown to occur also in the absence of such a strong lattice correspondence[15]. Electrostatic interactions between an ionic film headgroups and an ionic nucleating species may result in an arrangement of ions relative to a particular crystal face[16]. Despite the different mechanism through which epitaxial growth can

occur, most crystallisation experiments have shown that optimal conditions are achieved when the monolayer is compressed at high surface pressures, *i.e.* in a condensed, close-packed and fully ordered state[11,17,18] (Figure 1).

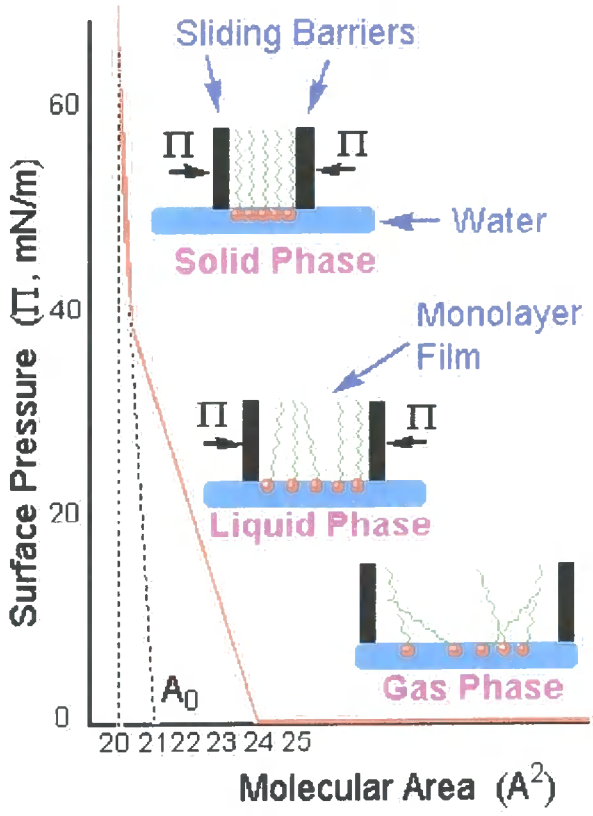


Figure 1 Typical Π-*A* curve and organisation of the surfactant molecules at different degrees of compression.

Thus, it has been long believed that a close-packed film is a necessary requirement for maximising the nucleation rate. The observed slower oriented growth beneath partially compressed or uncompressed monolayers has been so explained by assuming the existence in the expanded film of close-packed islands still able to act as nucleation promoters[10,19,20]. The presence of such ordered clusters has been in some cases demonstrated[21]. However, in monolayers with aliphatic chains, close-packed film islands are not generally observed at room temperature and low surface pressures. Moreover, films at lower surface pressure may also possess a certain degree of order, even in the absence of cluster formation. In 1996, Ahn and co-workers[22]

showed that calcite could crystallise beneath monolayers at a surface pressure equal to 10 mN m^{-1}, due to the capability of the organic template to reorganise itself in order to optimise the geometric and stereochemical match with the growing crystal face. In 1997, Cooper *et al.*[9, 23] showed for the first time that partially compressed or uncompressed films can work as preferential nucleation catalysts compared to their fully ordered counterparts. dl-Aspartic acid and d,l-asparagine were found to nucleate preferentially beneath monolayers at low to medium surface pressures (< 25 mN m^{-1}). Since the corresponding close-packed films were ineffective at inducing nucleation, the oriented growth observed at lower surface pressures was unlikely to be promoted by ordered clusters present in the film. Instead, the increased compressibility of these monolayers would allow them to reorganise to form arrangements which more closely resemble planes in the bulk crystals. For systems where there is not an exact match between the film and the nucleating species lattice structures, nucleation can be expected to be biased toward low to medium pressure, as the condensed films are more rigid and, consequently, less able to accommodate the existing mismatch than the corresponding partially compressed or uncompressed ones.

In this work, preliminary studies of the nucleation of dl-aspartic acid beneath monolayers constituted of the two protected glycopolymers, pAcGlcEMA and pAcGalEMA, were performed. Both films were found to act as efficient nucleation promoters, in a qualitatively and quantitatively similar manner.

Based on the knowledge that the nucleating ability of a film is paralleled with the capability of its soluble analogue to inhibit the growth of the same crystal face[24], the deprotected polymers, pGlcEMA and pGalEMA, were tested as dl-aspartic acid growth inhibitors. Glycoproteins are known to act as antifreeze in the serum of deep-sea fish, allowing them to survive at temperatures as low as -2 °C. Their property to

lower the freezing point seems not to be related to any colligative effect; on the contrary it is thought to arise from a mechanism that inhibits ice nucleation and growth[25-27]. Synthetic glycopeptide analogues have been prepared and shown to exhibit significant antifreeze properties[28,29].

3.2 RESULTS AND DISCUSSION

3.2.1 Π-*A* ISOTHERMS

Π-*A* isotherms for the protected polymers, pAcGlcEMA (**4a**), pAcGalEMA (**4b**) and pAcLacEMA (**10**), over aqueous solutions were recorded. Comparison of curves obtained for the same film over water and over a saturated dl-aspartic acid solution shows whether the amino acid adsorbs onto the film, an essential requirement for its heterogeneous nucleation beneath the monolayer.

Initially, measurements using octadecanoic acid, a film material of well-known Π-*A* behaviour, were carried out, in order to test the suitability of the apparatus used. Measurements were repeated changing different parameters until the curve obtained showed good agreement with the isotherm reported in the literature[30] and good reproducibility. Figure 2 shows that stable isotherms could be obtained. The limiting area per molecule was found to be ≈ 19.5 $Å^2$, in good agreement with the value of ≈ 20 $Å^2$ reported in the literature for isotherms recorded at 25 °C. It has to be noted that all the isotherms determined in this study were recorded at 20 ± 0.1 °C, unless differently stated, and with an accuracy of ± 1 mN m^{-1} and ± 1 $Å^2$. At Π = 0, the area per molecule of the film was ≈ 31 $Å^2$.

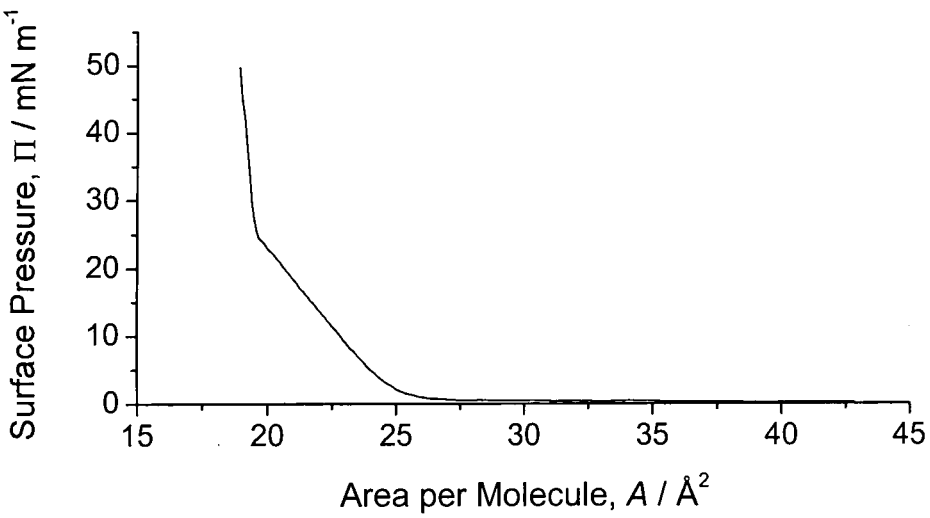


Figure 2 Π-*A* curve of octadecanoic acid on pure water at 20 °C.

The isotherms determined for all glycopolymers over pure water at neutral pH are reported in Figure 3.

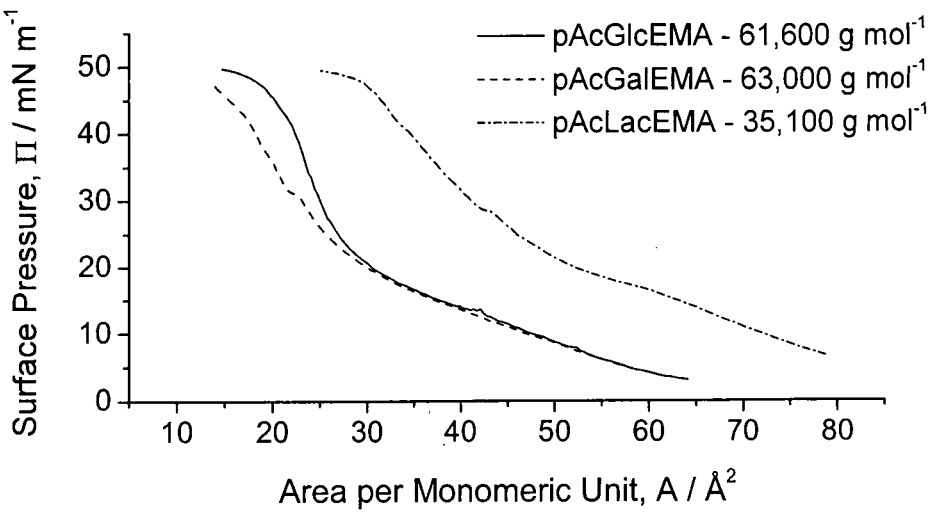


Figure 3 Π-*A* curves of pAcGlcEMA, pAcGalEMA and pAcLacEMA over pure water at neutral pH and at 20 °C. The number-average molecular weights of the polymers are indicated in the legend.

All glycopolymers were highly soluble in chloroform so that all solutions were easily spread on the subphase, as observed by eye and confirmed by optical microscopy. As

expected for polymer monolayers[31], the Π-*A* curves were smooth, without marked discontinuities indicating phase transitions. Limiting areas per monomeric unit at high surface pressure were ≈ 32 Å^2 for pAcGlcEMA, ≈ 33 Å^2 for pAcGalEMA and ≈ 60 Å^2 for pAcLacEMA. Limiting areas per repeating unit at Π = 0 were ≈ 71 Å^2 for **4a**, ≈ 72 Å^2 for **4b** and ≈ 100 Å^2 for **10**, as determined from experiments performed spreading a lower amount of material. Glycopolymers carrying monosaccharide residues (**4a** and **4b**) showed approximately the same limiting area (at high surface pressure) of ≈ 30 Å^2, despite the different structure of the peracetylated carbohydrate constituting the polar headgroup. For pAcLacEMA, instead, the limiting area per monomeric unit was double that determined for pAcGlcEMA and pAcGalEMA. In order to determine whether this result was related to the larger (double) dimension of the disaccharide residue or to the lower (half) molecular weight of the polymer, additional measurements were carried out using a different batch of pAcLacEMA having an average-molecular weight of 325,200 g mol^{-1}. Unfortunately, it was not possible to obtain **10** with M_n ≈ 60,000 g mol^{-1}. However, measurements performed using a polymer possessing a much larger molecular weight should indicate whether the increase of polymer dimension results in a condensation of the spreading film. The isotherms determined using the two different batches of pAcLacEMA are reported in Figure 4. Interestingly, the Π-*A* curve obtained for **10** having the highest molecular weight was slightly more expanded than that for the same polymer with the lowest M_n. Limiting areas per monomeric unit at high surface pressure varied from ≈ 60 Å^2 to ≈ 70 Å^2 with increasing the molecular weight of approximately 10-fold. Therefore, it could be concluded that the double limiting area per monomer unit obtained for pAcLacEMA, compared to the glycopolymers containing the monosaccharide residues, could be related to the double dimensions of the lactose residues, which

probably dispose themselves at the interface with the two sugar rings parallel to the surface.

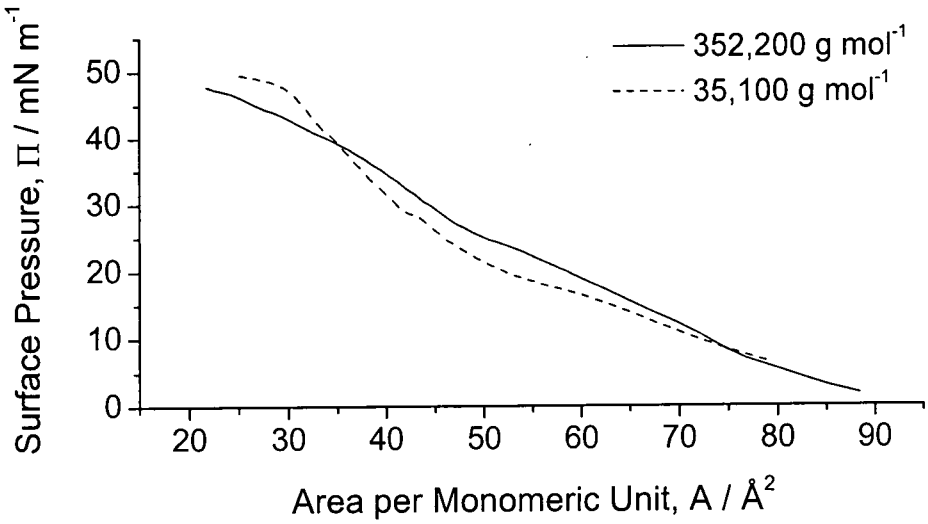


Figure 4 Π-*A* curves of pAcLacEMA over pure water at neutral pH and at 20 °C. The number-average molecular weights of the polymers are indicated in the legend.

Studies utilising polymer fractions of different average molecular weight as film materials have been reported[31] and the area occupied per monomer unit at surface pressures above 1 mN m^{-1} was generally found to be independent from the molecular weight.

In this study, further investigations of the effect of the molecular weight were carried out using two different batches of pAcGalEMA possessing different molecular weights. The isotherms so obtained are reported in Figure 5. In this case the results were rather different and the limiting area per monomeric unit increased with decreasing the molecular weight of the polymer. In particular, when the molecular weight roughly halved, the limiting area at high surface pressure increased from ≈ 33 $Å^2$ to ≈ 42 $Å^2$.

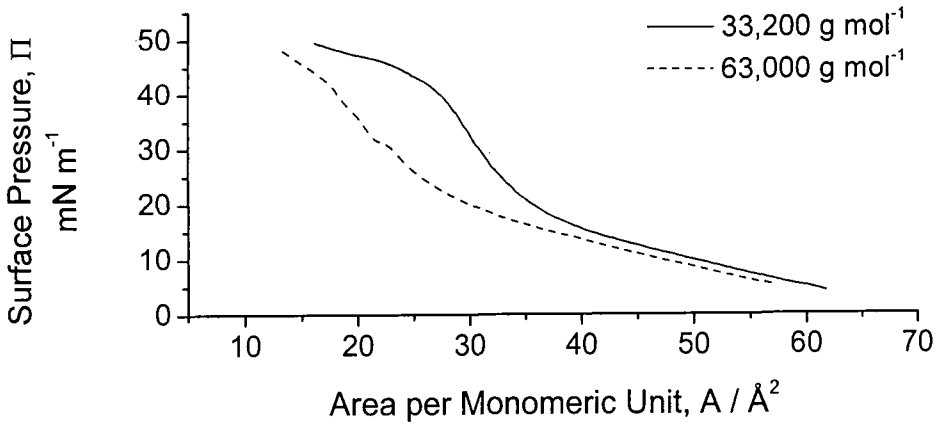


Figure 5 Π-*A* curves of pAcGalEMA over pure water at neutral pH and at 20 °C. The number-average molecular weights of the polymers are indicated in the legend.

After spreading and evaporation of the solvent, the freshly formed monolayer is usually not in a condition of absolutely stable equilibrium. Various processes producing film loss from the surface can then occur in order to reach this equilibrium (*e.g.* evaporation, solution, collapse or fracture)[32-34]. Therefore, the loss of material can be considered an index of the film stability. The Π-*A* curves of unstable films will progressively shift toward lower areas per molecule with repeating Π-*A* runs. If the presence of the interface provides a lower free energy than does the presence of an additional bulk phase, an extremely high insolubility and non-volatility of the film material can ensure that non-significant loss from the surface will take place. The monolayer state can so correspond to the minimum free energy of the system. However, even if the monolayer state does not represent the minimum of free energy, the film can still remain in its metastable state throughout the experiment, due to the slow rates of the processes producing material loss. Films of highly insoluble polymers are expected to be stable on the time-scale of the measurement, volatility not being a problem with macromolecules. Due to the decreased stability of the first formed aggregates of the nucleating species[35], in the absence of any suitable particles

on which the film can nucleate, many monolayers can be compressed to pressures considerably higher than their equilibrium spreading pressure – the pressure of the monolayer in equilibrium with its bulk phase – without collapse occurring. Eventually, however, there will be a surface pressure – collapse pressure, Π_c[36] – above which molecules will be forced out of the monolayer to form agglomerates of a new adjacent bulk phase. Either the area of the film will decrease rapidly at constant pressure or the pressure will drop if the film is maintained at constant area.

All glycopolymers formed highly stable films with considerably high collapse pressure: ≈ 40-50 mN m^{-1}. Moreover, in all cases repeated consecutive runs with increasing final surface pressure showed that the films could be compressed up to $\Pi \approx$ 35 mN m^{-1} before a shift of the Π-*A* curves toward lower areas per monomeric unit could be detected. All films were compressed at surface pressures 10-50 mN m^{-1}. The decrease of surface pressures after the films were held at Π values between 25 and 35 mN m^{-1} can be attributed to reversible reorganisation of the monolayers. The repeated runs recorded for pAcGlcEMA are reported in Figure 6.

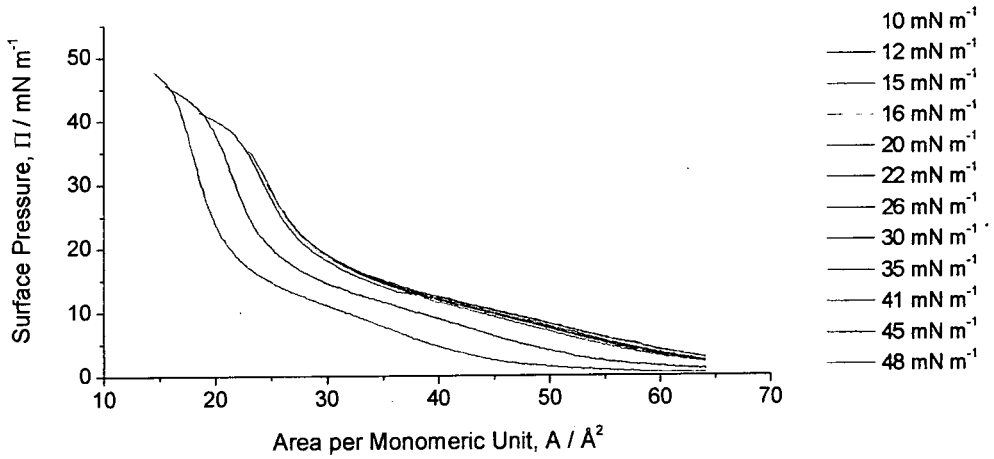


Figure 6 Π-*A* curves of pAcGlcEMA obtained performing consecutive runs at increasing final surface pressure, on pure water, at neutral pH, at 20 °C. The values of the final Π for each run are reported in the legend.

Saturated dl-aspartic acid solutions have pH ≈ 3. Π-*A* curves over these solutions should therefore be compared with the corresponding curves over water at the same pH, to negate any effect due to pH alone. Π-*A* curves over water at pH ≈ 3 were so determined. As expected, the isotherms using water at neutral pH and water at pH ≈ 3 as subphases did not show relevant differences. The curves obtained for pAcGlcEMA and pAcGalEMA are reported in Figure 7.

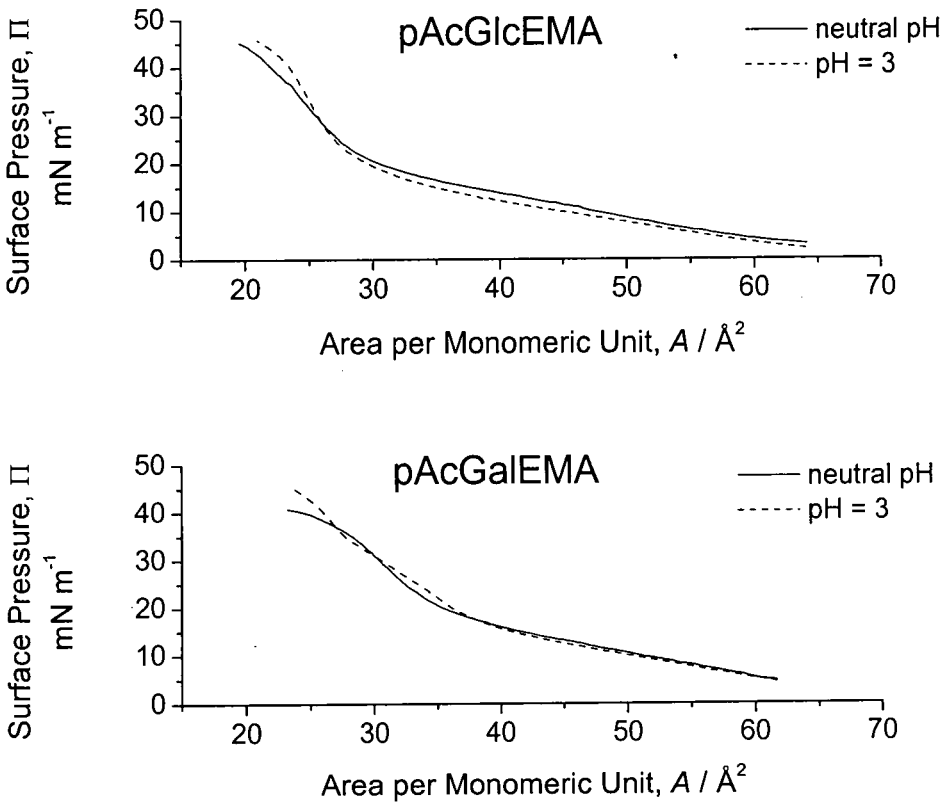


Figure 7 Π-*A* curves of pAcGlcEMA and pAcGalEMA on pure water at neutral pH and pH = 3, at 20 °C.

The presence of solute may have a profound effect on the film behaviour[37,38]. Its adsorption onto the film can occur when film-solute and water-water interactions exceed film-water and solute-water interactions. Thus, if the film-solute interactions

are sufficiently strong and the adsorption does not significantly reduce the favourable film-film interactions, the solute will strongly adsorb onto the film. As a consequence, the limiting area per molecule will be larger than the corresponding one in the absence of the solute, reflecting the cross-section of the film headgroup and the attached solute molecule, and the Π-*A* curve will shift toward the right[39]. The Π-*A* curves of pAcGlcEMA and pAcGalEMA on saturated dl-aspartic acid solution at 20 °C are reported in Figure 8.

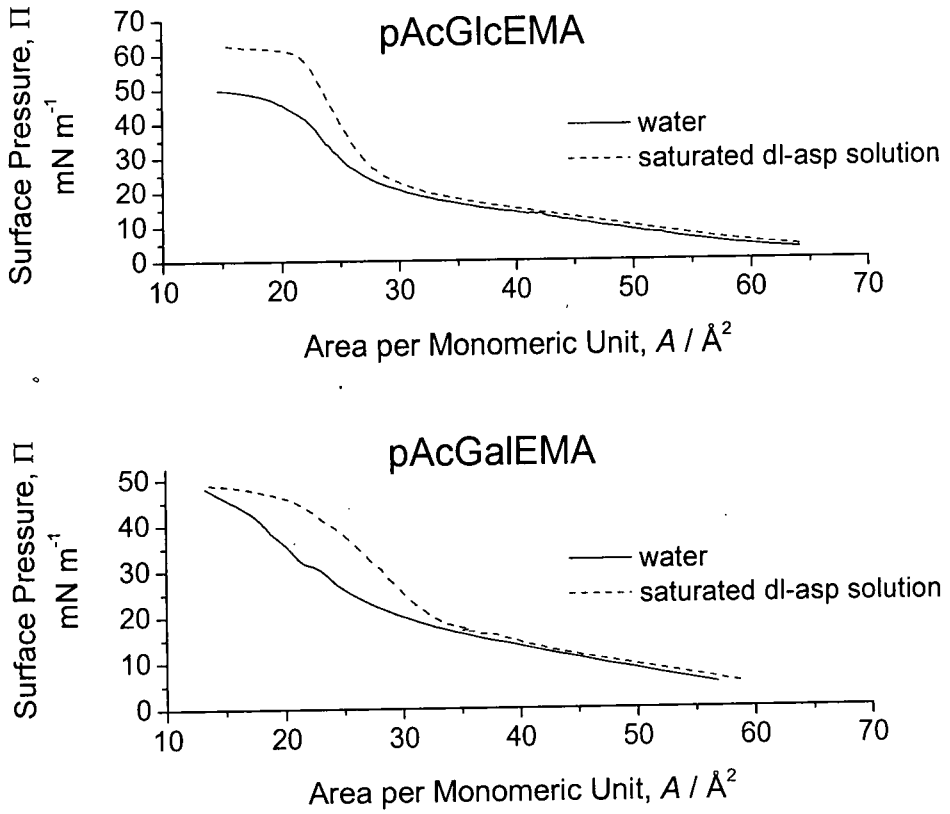


Figure 8 Π-*A* curves of pAcGlcEMA and pAcGalEMA on pure water and on saturated dl-aspartic acid solution at pH ≈ 3 and at 20 °C.

In both cases, adsorption of amino acid molecules between the film headgroups occurred at all surface pressures higher than approximately 15 mN m^{-1}, causing the observed slight expansion of the curves above this Π value. Adsorption was also

retained at high surface pressure, leading, for pAcGlcEMA, to a significant stabilisation of the film, as shown by the increase of the collapse pressure. Π-*A* curves above saturated dl-aspartic acid solutions were determined only for glycopolymers carrying the monosaccharide residues, since only these materials were then tested as nucleation promoters. As already mentioned in Section 3.1, adsorption at the surface of the nucleating species is the first requirement for its film induced epitaxial growth.

3.2.2 NUCLEATION OF DL-ASPARTIC ACID BENEATH GLYCOPOLYMER LANGMUIR FILMS

The effect of pAcGlcEMA and pAcGalEMA films on dl-aspartic acid nucleation was investigated, using subphase amino acid solutions with different percentage supersaturation, σ%, and working at different surface pressures, *i.e.* at different degrees of film compression. Supersaturated solutions of the amino acid were prepared using the equation and data reported in Appendix II.

Initially, pAcGalEMA (M_n = 63,000 g mol^{-1}) was chosen as the film material and a dl-aspartic acid solution with σ% = 120% as the subphase. The crystallisation experiments were carried out at 10 °C, in order to minimise the water evaporation from the trough throughout the experiment (24 h). In the first experiment the film was compressed at a surface pressure of 25 mN m^{-1} and the temperature kept at 10 ± 0.1 °C by means of a water bath connected to the trough. The temperature of the subphase surface was checked by means of a thermometer equipped with a surface probe. A temperature gradient of 3 degrees was observed between the solution surface and the PTFE base of the trough, with the temperature of the latter being 7 °C. This caused an increase of the supersaturation of the solution in the proximity of the trough base, in turn responsible for the observed significant nucleation and growth of crystals on

scratches present on the bottom of the apparatus. Figure 9 reports a picture taken from the base of the trough. It can be seen that the crystals showed the morphology characteristic of crystals grown homogeneously from aqueous solutions, with a prismatic crystal habit, dominated by low energy {110} faces, edged by two $\{\bar{1}11\}$ and two $\{\bar{2}02\}$ faces (see Appendix III). All faces were smooth.

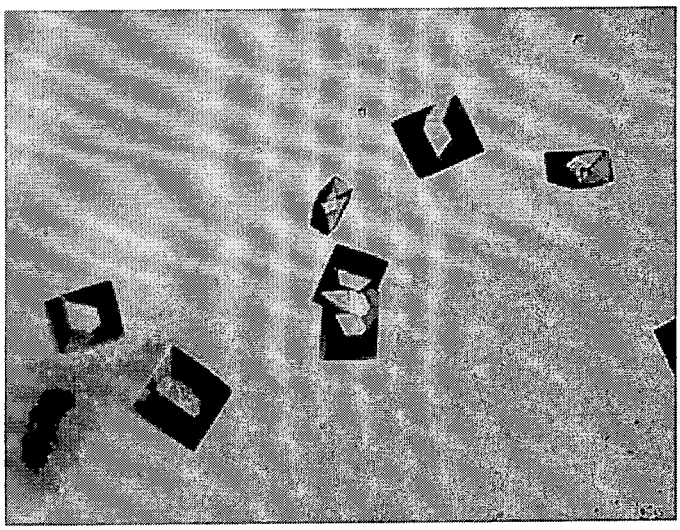
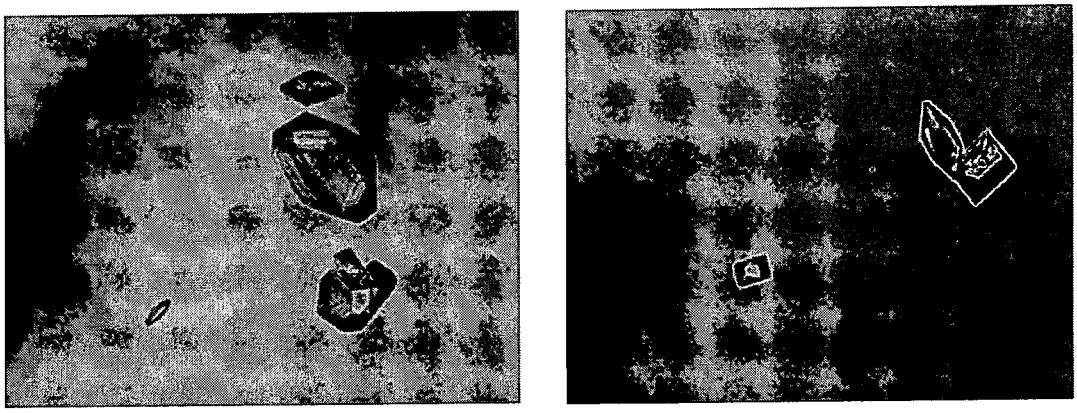


Figure 9 Typical crystal habit of dl-aspartic acid nucleated homogeneously from aqueous solution. The picture was taken from the bottom of the trough after 24 h (magnify ×100).

Pictures of the crystals observed at the surface are reported in Fgure 10. Crystals grown beneath the pAcGalEMA monolayer film presented a different crystal habit with respect to those grown on the base of the trough, with an enlargement of the face grown attached to the monolayer. Crystallisation beneath monolayer films occurs, after adsorption of the nucleating species, by crystal nucleation at the surface followed, after attainment of the nuclei critical size, by mainly diffusion limited crystal growth. Diffusion of crystallising material to the attached faces will be reduced compared to that toward the faces in solution, decreasing their growth rate and so resulting in their enlargement (see Figure 21). Moreover, faces grown attached to the film showed numerous surface irregularities, whilst all other faces appeared smooth.

A large number of surface structure effects is typical of epitaxial growth, due to either coalescence of growing nuclei or to lattice mismatch between the film and the nucleating crystal face. Since the nucleation rate will cause less than 1% of the surface to be covered by crystals, in the present system irregularities can be assumed to result predominantly from misfit dislocations due to the inherent lattice mismatch between the film and the nucleating species.



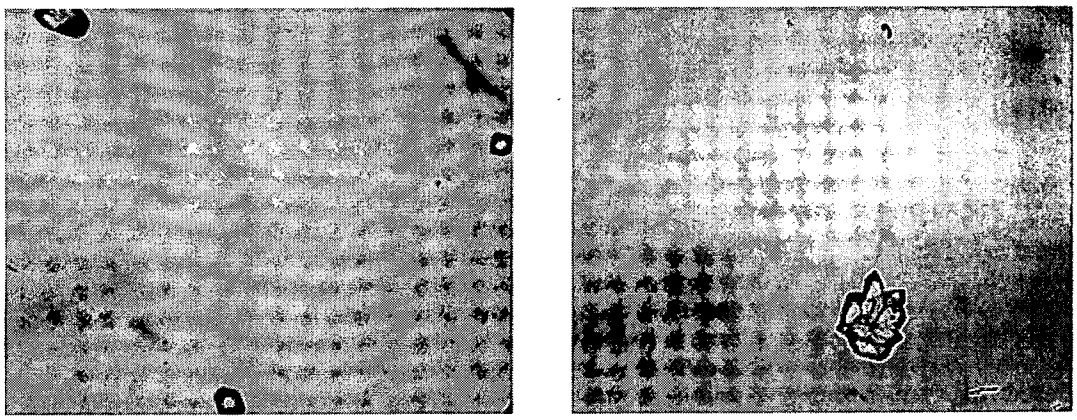
(a) (b)

Figure 10 dl-Aspartic acid crystals grown after 24 h beneath pAcGalEMA film at Π = 25 mN m^{-1} from a 120% supersaturated solution at 10 °C (magnify ×100). a), Faces $\{\bar{1}\,11\}$ and {110} grown attached to the film; b), {110} and probably {200} faces grown bounded to the film.

In order to avoid the previously mentioned gradient of temperature, critical for this kind of investigation, it was decided to carry out the experiments at 20 °C. Nucleation on the base and sides of the PTFE trough, resulting in a lowering of the supersaturation level throughout the experiment, should, in fact, be avoided. The temperature was recorded for the initial and final instant of the experiments and it was found to remain constantly equal to 20 °C. However, water evaporation from the trough during the 24 h, with subsequent increasing of the dl-aspartic acid solution supersaturation, was found causing the loss of the solution up to *ca.* 30% of the total volume. This undesired phenomenon was afterwards reduced by placing two 2 litres

beakers of water (kept at 70 °C throughout the experiment) under the sheet covering the trough. The loss of solution for each experiment was then *ca.* 15% of the total volume. The evaporation of the water from the trough also causes an error in the value of the surface pressure measured. Due to the decrease of the depth of immersion of the Wilhelmy plate during the experiment, the value of Π will decrease with time. However, the magnitude of this effect was found to be negligible compared to lowering in the surface pressure due to reorganisation, dissolution, fracture and collapse of the film. Despite the fact that a supersaturation of 120% is well below the reported value of 284 ± 5%[9] as the critical supersaturation limit for homogeneous nucleation of dl-aspartic acid, again crystallisation upon scratches present on the base of the trough was found to invalidate the measurements. It was, therefore, decided to reduce the α% of the subphase to 43%. pAcGalEMA (M_n = 63,000 g mol^{-1}) was again used as the film material. Two experiments were carried out, at 10.5 mN m^{-1} and 20 mN m^{-1}, respectively. As expected practically no crystals were observed in the first case, due to negligible adsorption of amino acid molecules onto the film at this low surface pressure (see Figure 8). However, crystallisation was poor also for the highest surface pressure, where the isotherm reveals that adsorption of the nucleating species is occurring. Pictures of crystals grown at the surface inside the barriers in these conditions are reported in Figure 11. It is worth noting that at this low supersaturation, nucleation of dl-aspartic acid onto the base of the trough was negligible and the few crystals observed (grown onto the trough base) showed the typical habit of crystals nucleated homogeneously (see Figure 9). The film was stable throughout the experiment, the value of the surface pressure recorded after 24 h being equal to 17

mN m^{-1}. Crystal faces nucleated attached to the monolayer appeared to be mainly {110} and $\{\bar{1}\,11\}$*.



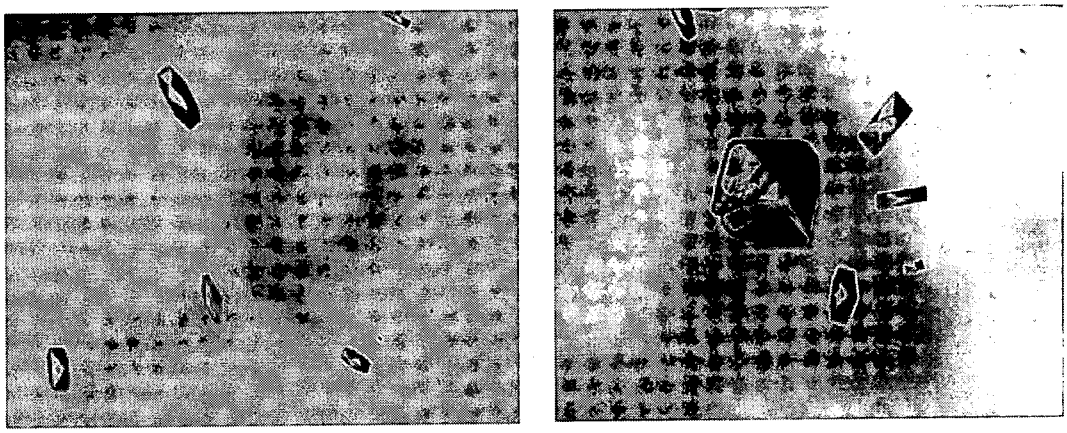
(a) (b)

Figure 11 dl-Aspartic acid crystals grown after 24 h beneath pAcGalEMA film at Π = 20 mN m^{-1} from a 43% supersaturated solution at 20 °C (magnify ×100). a), Small crystals; b), clump.

In order to determine whether the scarce crystallisation was due to the low supersaturation of the solution used as the subphase or to an inherent poor capability of the film to act as a nucleation catalyst, further experiments were carried out using dl-aspartic acid solution with α% equal to 100%. pAcGalEMA (M_n = 63,000 g mol^{-1}) was again employed as the film material and the monolayer compressed at 15 mN m^{-1}, the limit at which the amino acid was shown to adsorb onto the film. After 24 h, the number of crystals observed at the surface between the barriers was significantly increased compared to the previous experiments, but the crystals were still generally small. No crystals were observed at the surface outside the barriers and in the supersaturated solution kept as a control. Very few crystals were detected on the base of the trough. Again these latter crystals presented the habit typical of homogeneous nucleation reported in Figure 9. Hence, any difference in the morphology of crystals

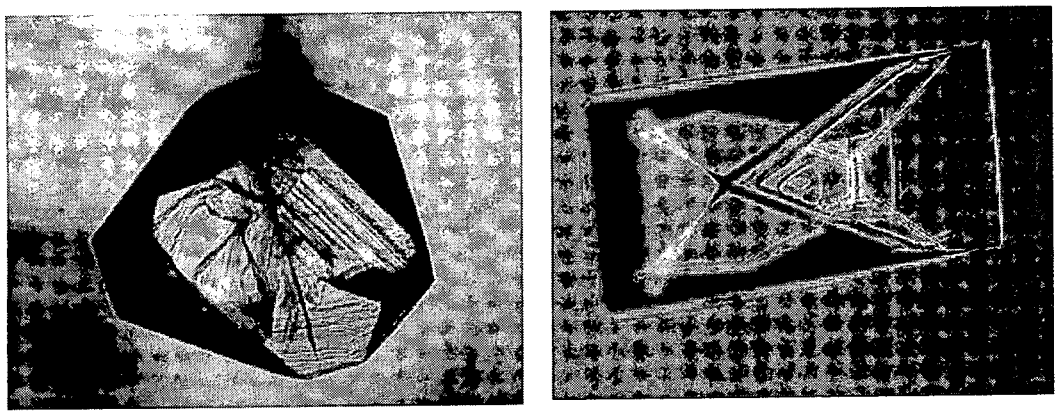
* See Appendix III for the assignment of the crystal faces.

grown heterogeneously under the monolayer could be attributed to the film and not to any supersaturation effect. The film was found to be rather stable during the experiment, with a final surface pressure equal to 7 mN m^{-1}. Pictures of crystals grown in these conditions are reported in Figures 12 and 13.



(a) (b)

Figure 12 a-b), dl-Aspartic acid crystals grown after 24 h beneath pAcGalEMA film at Π = 15 mN m^{-1} from a 100% supersaturated solution at 20 °C (magnify ×100).



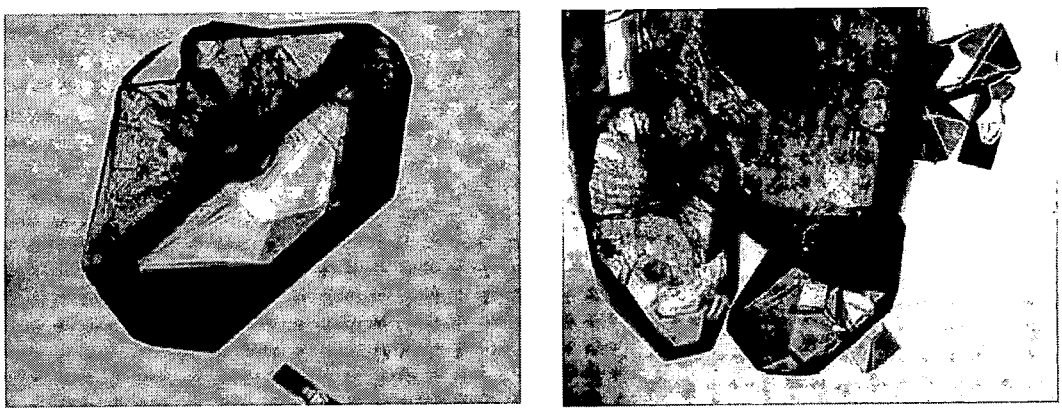
(a) (b)

Figure 13 dl-Aspartic acid crystals grown after 24 h beneath pAcGalEMA film at Π = 15 mN m^{-1} from a 100% supersaturated solution at 20 °C (magnify ×500). a), $\{\bar{1}\,11\}$ Face grown attached to the film; b), {110} face grown bounded to the film.

Faces grown attached to the film were typically enlarged compared to their usual size in crystals grown from aqueous solution. These faces were mainly {110} and $\{\bar{1}\,11\}$.

They showed a high density of surface irregularities, with growth appearing to occur outwards in all directions from a site of initial nucleation ("hopper" growth[40], Figure 13b). All other faces present in the crystals were those commonly seen in aqueous solution, although the large size of the faces attached to the film would sometimes mask the formation of the adjacent crystal faces.

The experiment was left over the weekend and the crystals at the surface were observed again after two days (that is after three days from the beginning of the experiment). The size was significantly increased and numerous aggregates of large crystals could be observed, as reported in Figure 14.



(a) (b)

Figure 14 dl-Aspartic acid crystals grown after 72 h beneath pAcGalEMA film at Π = 15 mN m^{-1} from a 100% supersaturated solution at 20 °C (magnify ×100). a), Face $\{\bar{1}\,11\}$ grown attached to the film; b), aggregate.

The 100% supersaturated solution was used as the subphase for subsequent experiments, where pAcGlcEMA (M_n = 61,600 g mol^{-1}) and pAcGalEMA (M_n = 63,000 g mol^{-1}) were used as film materials and the monolayers compressed at 20 mN m^{-1}. In both cases the monolayers showed good stability throughout the experiments, with final surface pressures equal to 13 and 11 mN m^{-1} for pAcGlcEMA and pAcGalEMA, respectively. The evaporation of the solution was approximately 15% of the total volume. Both glycopolymer films were found to be effective dl-aspartic

acid nucleators. Numerous large crystals were well visible to the naked eye after 24 h, forming aggregates covering almost all the surface between the barriers. In both experiments no crystals were detected at the surface outside the barriers and in the supersaturated solution kept as a control. Some of the crystals nucleated heterogeneously were collected and kept for X-ray investigations that could be performed in the near future. Pictures of crystals obtained in these conditions are reported in Figures 15, 16 and 17.

(a) (b)

Figure 15 dl-Aspartic acid crystals grown after 24 h beneath pAcGlcEMA film at Π = 20 mN m^{-1} from a 100% supersaturated solution at 20 °C (magnify ×100). a-b), Aggregates with $\{\bar{1}\,11\}$, {110} and $\{\bar{2}\,02\}$ as attached faces.

(a) (b)

Figure 16 dl-Aspartic acid crystals grown after 24 h beneath pAcGalEMA film at Π = 20 mN m^{-1} from a 100% supersaturated solution at 20 °C (magnify ×100). a-b), Aggregates with $\{\bar{1}\,11\}$, {110} and $\{\bar{2}\,02\}$ as attached faces.

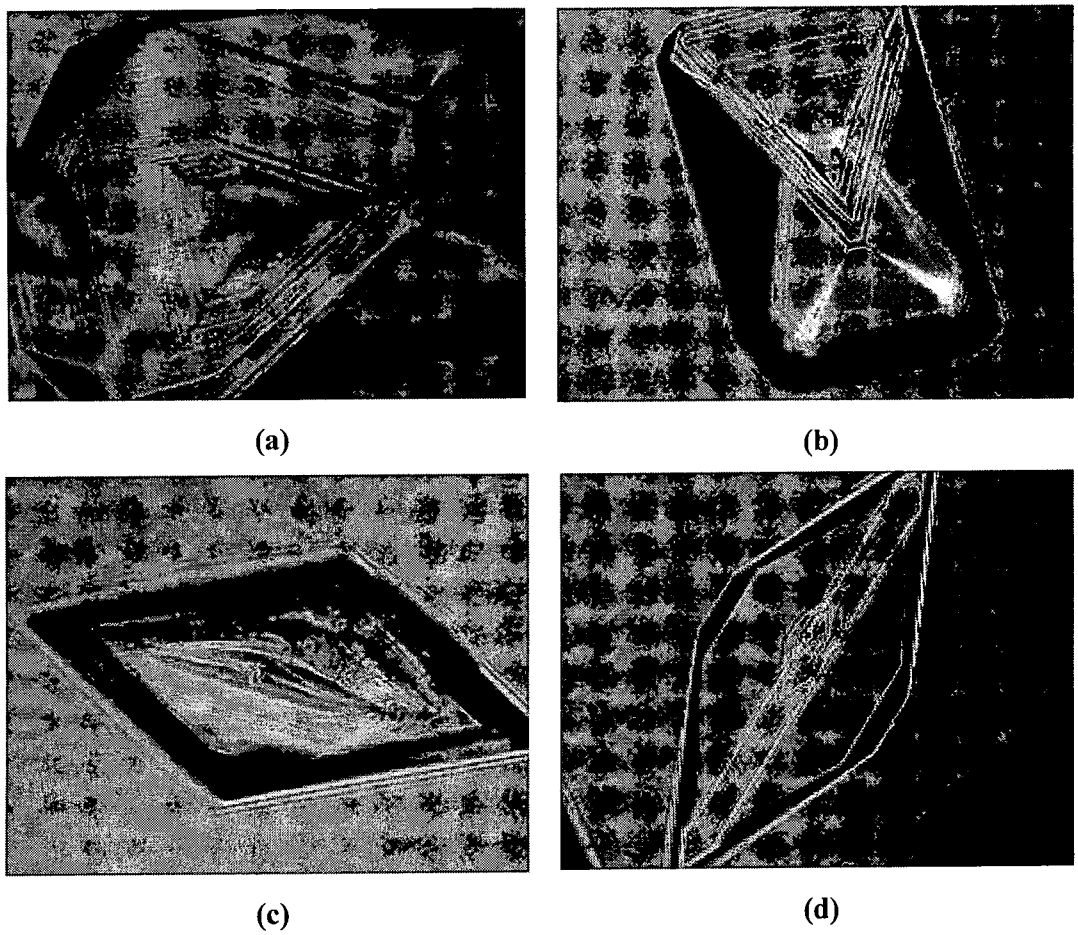


Figure 17 dl-Aspartic acid crystals grown after 24 h beneath pAcGlcEMA and pAcGalEMA films at Π = 20 mN m^{-1} from a 100% supersaturated solution at 20 °C (magnify ×500). a), $\{\bar{1}11\}$ Attached face (pAcGalEMA film); b), {110} attached face (pAcGlcEMA film); c), $\{\bar{2}02\}$ attached face (pAcGlcEMA film); d), $\{\bar{2}02\}$ attached face (pAcGalEMA film).

It can be seen that the crystal faces that grew attached to pAcGlcEMA and pAcGalEMA films in these conditions were namely $\{\bar{1}11\}$ and {110}, with a smaller proportion of $\{\bar{2}02\}$. In both cases, the largest crystals seemed to grow with faces $\{\bar{1}11\}$ bound to the films. Again the attached faces could be recognized by the high density of surface structure effects. The typical "hopper" growth, occurring outwards in all directions from a focal region, was also evident (Figure 17).

Further crystallisation experiments were carried out using the 100% supersaturated solution of dl-aspartic acid as the subphase and pAcGalEMA (M_n =

33,200 g mol^{-1}) as the film material. The limiting area per monomeric unit at high surface pressure for this polymer of lower molecular weight was ≈ 42 $Å^2$. The monolayer was compressed at 20 mN m^{-1}. Interestingly, the film was still found to act as a nucleation promoter, but its effectiveness was lower than that of the corresponding polymer with higher molecular weight with a limiting area per monomeric unit of ≈ 30 $Å^2$. Pictures of the crystals observed at the surface between the barriers are reported in Figure 18.

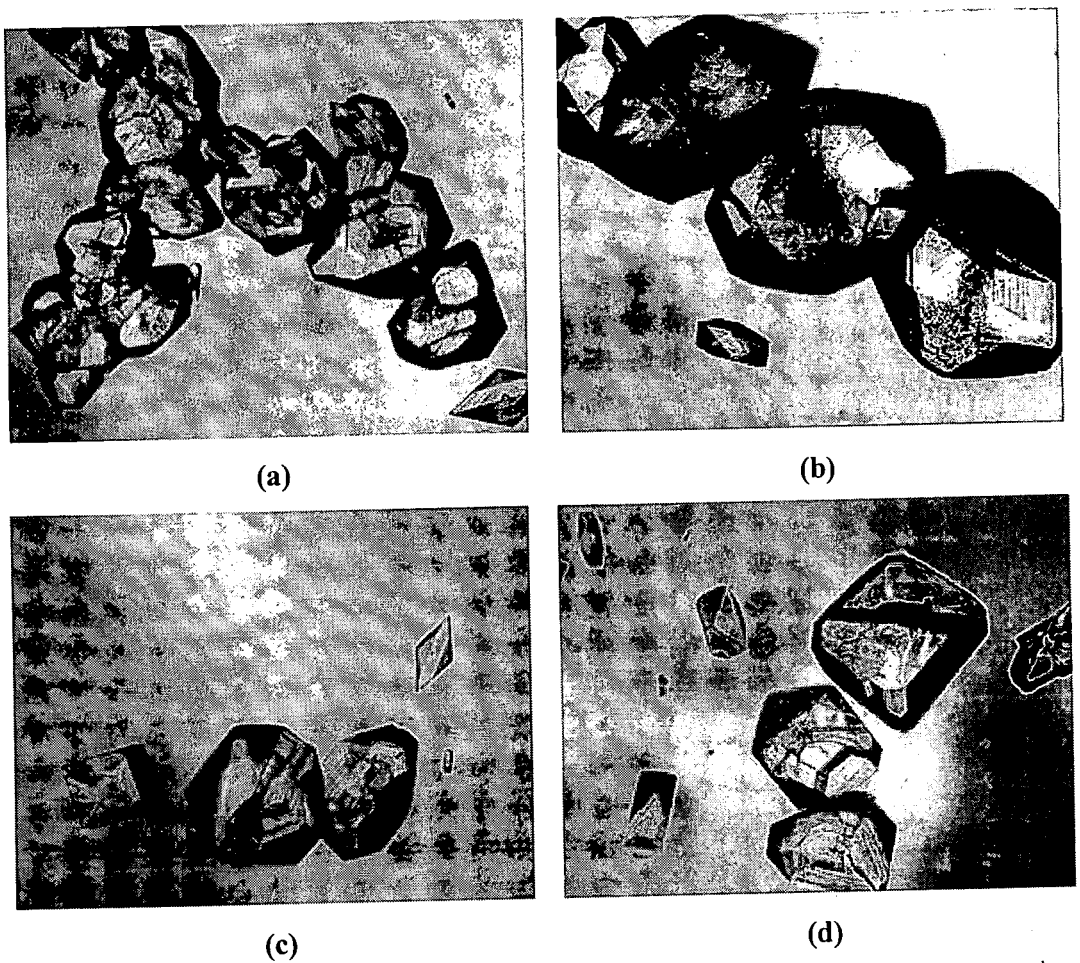


Figure 18 dl-Aspartic acid crystals grown after 24 h beneath pAcGalEMA (M_n = 33,200 g mol^{-1}) film at Π = 20 mN m^{-1} from a 100% supersaturated solution at 20 °C (magnify ×100). a-b), Aggregates; c-d), namely { $\bar{1}$ 11} faces grown attached to the film.

The crystals observed on the surface between the barriers after 24 h showed a morphology analogous to that determined from the previous experiments, with the $\{\bar{1}11\}$, $\{110\}$ and, in small quantity, $\{\bar{2}02\}$ faces grown attached to the film. Again, the largest crystals seemed to grow bounded through the $\{\bar{1}11\}$ faces. However, the number of crystals was significantly smaller than that reported for pAcGlcEMA (M_n = 61,600 g mol^{-1}) and pAcGalEMA (M_n = 63,000 g mol^{-1}), with few aggregates visible. Moreover, the crystals appeared to be larger and slightly more uniform in size than those showed in Figures 15-17. This suggests that once the crystal nuclei were formed, subsequent crystal growth occurred on these nuclei at the expense of fresh nucleation[17]. Isotherms recorded for this polymer over aqueous and saturated dl-aspartic acid solution (Figure 19) showed that the extent of amino acid adsorption onto the film at 20 mN m^{-1} was comparable to that determined for pAcGlcEMA (M_n = 61,600 g mol^{-1}) and pAcGalEMA (M_n = 63,000 g mol^{-1}) at the same surface pressure.

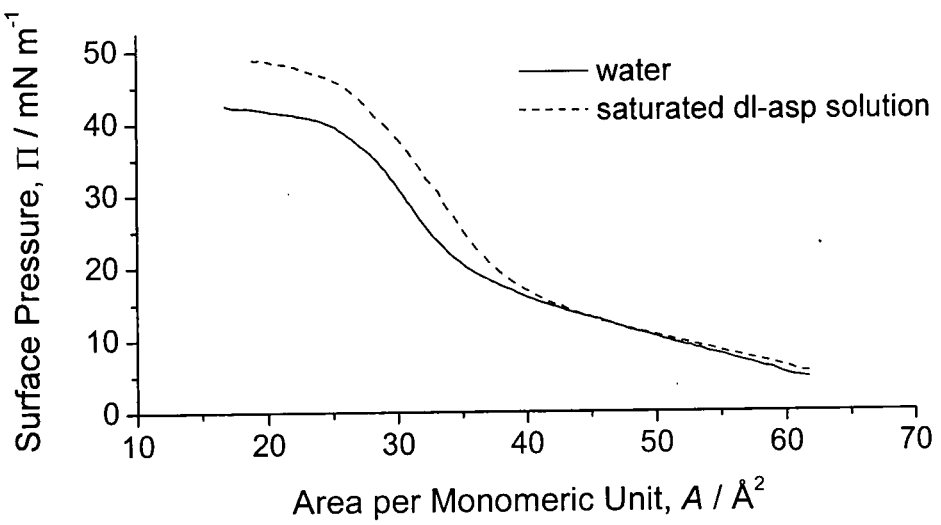


Figure 19 Π-*A* curves of pAcGalEMA (M_n = 33,200 g mol^{-1}) on pure water and on saturated dl-aspartic acid solution at pH ≈ 3 and at 20 °C.

Since, as has already been pointed out, nucleation beneath monolayers relies on a degree of molecular recognition between the film and the nucleating species, the structural/geometric correlation must be lower for pAcGalEMA (M_n = 33,200 g mol^{-1}) than for pAcGlcEMA (M_n = 61,600 g mol^{-1}) and for pAcGalEMA (M_n = 63,000 g mol^{-1}). In other words, the larger area per monomeric unit of the former Langmuir monolayer may result in a greater mismatch between the film and the growing crystal lattice structures; increased mismatch that could not be fully accommodated by the very likely higher compressibility of this film compared to that of pAcGlcEMA and of pAcGalEMA with the highest molecular weight. Thus, whilst a significant adsorption of the nucleating species could still take place onto this more expanded film, the resultant arrangement of dl-aspartic acid molecules was less suitable to induce crystallisation of the bulk phase, decreasing the density of crystals obtained.

More experiments would be required in order to assign rigorously the mechanism through which these glycopolymer films induce epitaxial growth of dl-aspartic acid. Experiments at higher surface pressures would reveal if nucleation would be favoured by increasing the order of the organic template. Most importantly, investigations dedicated to the determination of the structure of the polymeric films would be essential. Polymers, in fact, behave rather differently from monomeric surfactants. It is already difficult to decide whether or not polymers are forming true monolayer films. Thus, for example, it is generally assumed that the polymer disposes itself at the surface so that each monomer segment is constrained to remain anchored at the interface throughout the experiment. However, it is easily possible that each polymer molecule has only a few of its polar headgroups anchored to the surface, while the rest actually extends into the adjacent bulk phase[31]. It is also possible that the polymer maintains at the surface its solution structural organisation (*i.e.* coils), as

has been found for some proteins for which the film is viewed as formed from intact α-helices[41].

Nevertheless, keeping in mind that there is no straightforward relation between the results obtained for monomeric surfactants and polymeric materials, a latter aspect of the results obtained in the present study may be underlined. It has been found that when epitaxial growth occurs under a partially compressed or uncompressed film, thanks to the capability of the latter to reorganise itself in order to optimise the geometric and stereochmical match with the to-be-grown crystal, crystallisation is less specific, leading to more than one crystal morphology. The film is able to reorganise into more arrangements that mimic planes in the bulk crystal and different crystal faces can nucleate at the same surface pressure[9]. This behaviour was observed in this study, where crystals were found to grow with faces $\{\bar{1}11\}$, $\{110\}$ and, in small quantity, $\{\bar{2}02\}$ attached to the films. This would, therefore, suggest that pAcGlcEMA and pAcGalEMA films possess a certain degree of compressibility which allow them to accommodate eventual mismatches with the nucleating species, leading to arrangements of the adsorbed dl-aspartic acid molecules compatible with the $\{\bar{1}11\}$, $\{110\}$ and $\{\bar{2}02\}$ faces.

3.2.3 Inhibition of dl-Aspartic Acid Growth by Glycopolymers

Preliminary studies of the crystal growth inhibition activity of the deprotected analogues, pGlcEMA (**5a**) and pGalEMA (**5b**), were performed. Crystals were grown at room temperature in a 100% supersaturated dl-aspartic acid solution at 20 °C, containing pGlcEMA or pGalEMA (Figure 20b-f). The habit of the crystals obtained was compared with that of crystals grown in supersaturated solution only (Figure

20a). The presence of the polymers in solution seemed to affect the morphology of the resulting crystals.

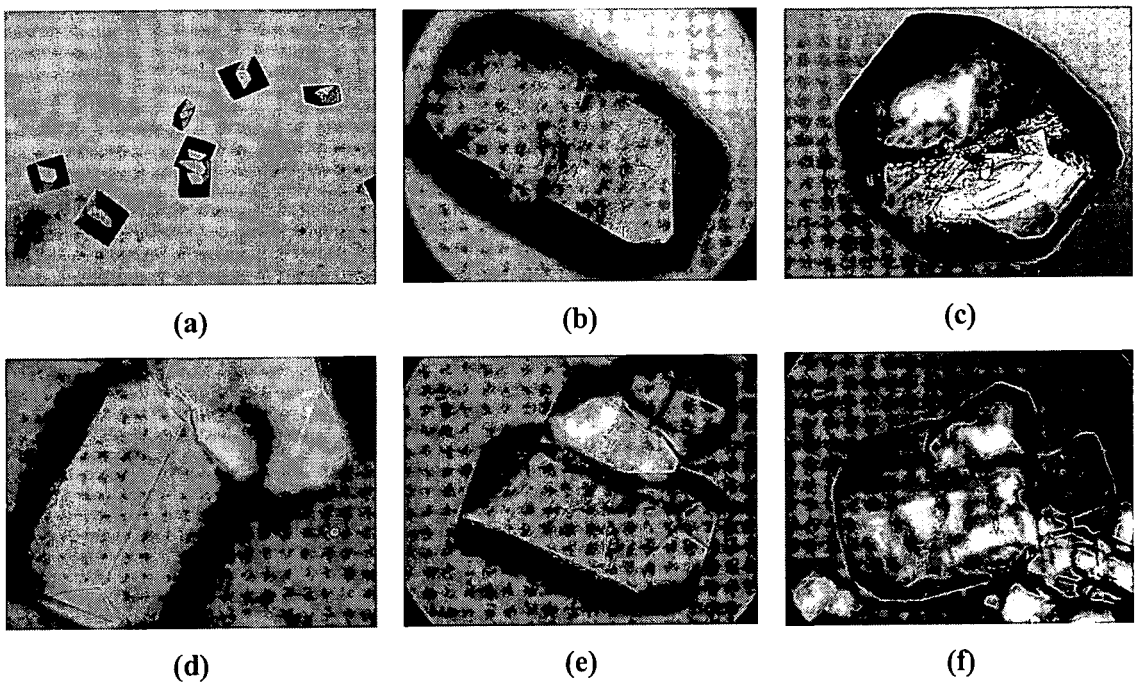


Figure 20 a), dl-Aspartic acid crystals grown in homogeneous conditions in supersaturated solution only; b-c), dl-aspartic acid crystals grown in 100% supersaturated solution containing 10 mg ml^{-1} of pGalEMA; d-f), dl-aspartic acid crystals grown in 100% supersaturated solution containing 10 mg ml^{-1} of pGlcEMA.

Factors governing nucleation promotion and crystal growth inhibition are the same: both phenomena depend on the extent to which a foreign matrix can produce adsorption of, or adsorb onto, respectively, a particular crystal face. In the case of crystal growth inhibition the foreign material adsorbs onto a crystal face exposing its hydrocarbon "tails" at the surface and so reducing subsequent growth upon that face (Figure 21). Therefore, this face, having the lowest rate of growth, will dominate the resultant crystal morphology.

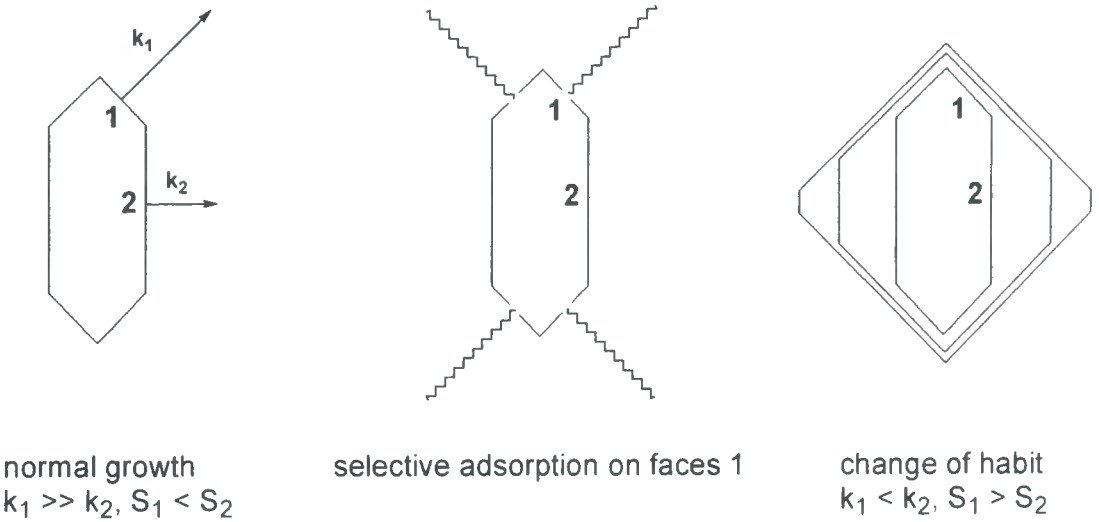
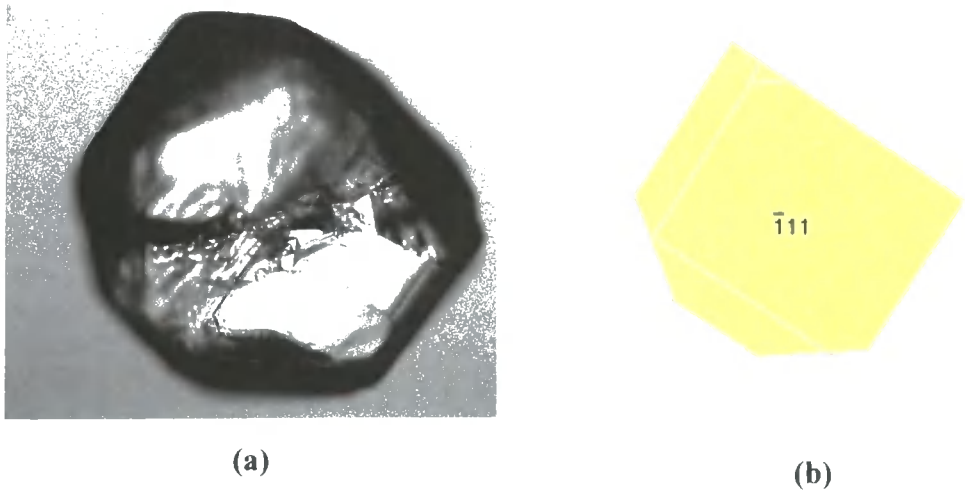


Figure 21 Crystal growth inhibition.

Comparison of the morphologies of the obtained crystals with those predicted by computational modelling[†], revealed that pGalEMA seemed to affect dl-aspartic acid growth by adsorbing onto faces $\{\bar{1}11\}$ (Figure 22) and {110}. pGlcEMA appeared to behave similarly and the majority of crystals obtained seemed to have {110} as the dominant face (Figure 23).

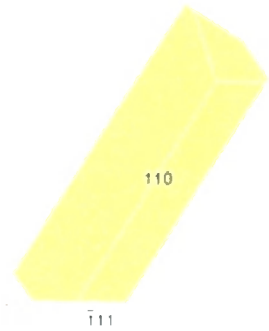


(a) **(b)**

Figure 22 a), dl-Aspartic acid grown in 100% supersaturated solution containing 10 mg ml^{-1} of pGalEMA; b), crystal morphology as predicted by computational modelling with Cerius^2, showing a dominant $\{\bar{1}11\}$ face and compatible with the morphology experimentally obtained.

[†] Cerius^2 program using a Bravais Friedel Donnay Harker (BFDH) method. The comparison was carried out by measurement of the face angles.

(a)



(b)

Figure 23 a), dl-Aspartic acid grown in 100% supersaturated solution containing 10 mg ml^{-1} of pGlcEMA; b), crystal morphology as predicted by computational modelling with Cerius^2, showing a dominant {110} face and compatible with the morphology experimentally obtained.

Although further investigations are absolutely required in order to assess whether these glycopolymers can actually work as crystal growth inhibitors, it is interesting to note that the pGlcEMA and pGalEMA appeared to inhibit the growth of faces $\{\bar{1}11\}$ and {110}, which are exactly those that predominantly nucleated beneath films of the protected, insoluble analogues. This correspondence suggests the reliability of the results obtained for pGlcEMA and pGalEMA.

3.3 EXPERIMENTAL

3.3.1 GENERAL

dl-Aspartic acid was purchased in its purest available form, 99%, from Sigma. The polymers used were prepared according to the procedures reported in Section 2.3.3. The octadecanoic acid used in the preliminary Π-*A* runs was 99% pure, supplied by Aldrich. Ultra-pure Milli-Q water with a resistivity of 18 MΩm was used throughout the experiments, both in preparing the dl-aspartic acid solutions and in all

the Π-*A* runs. Each film material was spread from a 99.9% HPLC grade chloroform solution, purchased from Aldrich.

3.3.2 TROUGH CLEANING PROCEDURE

Purity and protection from contamination are of utmost importance in monolayer experiments. Due to the very small amount of material used – typically properties of a film containing << 1 mole have to be measured –, any contaminant present can cause serious effects. Therefore, great care has to be taken to ensure that the troughs and the barriers are clean and the chemicals used are as pure as possible. A scrupulous trough cleaning procedure was adopted: before each experiment, the trough and the barriers were first rinsed with ultra-pure Milli-Q water. Then, they were cleaned with a special surfactant-free tissue soaked in chloroform and again rinsed with ultra-pure Milli-Q water. During all the procedure disposable powder-free plastic gloves were worn to prevent grease contamination from the hands.

3.3.3 CRITERIA FOR A CLEAN SUBPHASE SURFACE

Before spreading of the film on the subphase, the surface of the latter was checked for any surface impurities. The trough was filled with the subphase solution with the barriers at their maximum area position. The surface pressure of the subphase was taken as zero. The barriers were then moved inwards until their minimum area position. When the balance reading did not increase by more than 0.5 mN m^{-1} during the compression, the subphase solution was considered suitably free from surface active contaminants. Any surface impurity present could be removed from the surface using a clean pipette attached to a water pump until a clean surface was obtained.

3.3.4 FILM SPREADING

Ultra-pure chloroform was chosen as a suitable solvent for the spreading of the film materials. The films were added drop-wise to the subphase from ≈ 1 mg ml^{-1} solutions, using a micro-syringe, which could measure 1 μl quantities of the film solution. The syringe was held no further away than 1 cm from the surface to prevent any film solution sinking to the trough bottom. The solvent was allowed to evaporate between each addition by waiting 10 seconds before the next drop was added.

3.3.5 Π-*A* ISOTHERMS

All Π-*A* isotherms were recorded using an automated Langmuir trough (NIMA Technology Ltd.), type 601A (30 × 20 × 0.5 cm^3), containing ≈ 350 cm^3 when filled so that a convex meniscus formed (Figure 24).

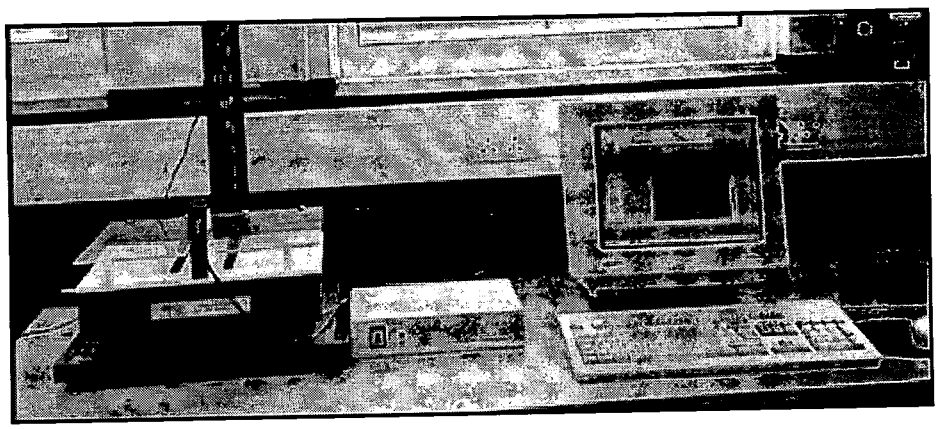


Figure 24 Langmuir trough (type 601A) used for Π-*A* curves determination.

The Wilhelmy plate was made of a piece of filter paper connected, via a light metal hook, to a microbalance. If it is assumed that the contact angle, θ, between the Wilhelmy plate and the liquid phase in the trough is zero, *i.e.* complete wetting of the

plate occurs, then it is possible to relate the weight measured by the microbalance to the surface pressure. A fresh plate was used for each experiment.

Typically, 50 μl of film solution of approximately 1 mg ml^{-1} were added to the surface. The film solution concentrations used were 1.00 mg ml^{-1} for octadecanoic acid, 1.04 mg ml^{-1} for pAcGlcEMA, 1.08 mg ml^{-1} for pAcGalMEA and 1.05 mg ml^{-1} for pAcLacEMA. After the spreading, the chloroform was allowed to evaporate waiting 3 minutes before the initial surface pressure reading was taken. The films were spread on pure water (neutral and acidic pH) and on saturated dl-aspartic acid solution. The barriers were then moved under computer control, changing the surface pressure (Π) and the area per repeating unit (A) of the films. Π-A isotherms were recorded protecting the trough with its transparent cover to avoid contamination of the surface.

3.3.6 PREPARATION OF SUPERSATURATED SOLUTIONS

The required amount (see Appendix II) of dl-aspartic acid was added to 500 ml of Milli-Q water and heated until all the crystalline material had dissolved. The solution was then filtered in a conical flask using the pre-heated filtering apparatus shown in Figure 25. The filtration was conducted under atmospheric pressure, since raising the pressure beyond this could cause the tearing of the 0.22 μm filters through which the solution flowed. Filters of this pore size removed any dust particles possibly present within the solution. After the filtration, the solution was kept, covered, at a temperature of at least 10 °C above its saturation temperature for two hours, in order to dissolve any microscopic crystal nuclei that could have passed through the filter. Then, the solution was left at room temperature overnight. The next day it was used for the nucleation studies, unless it contained crystalline material. The

remainder of the solution was kept covered throughout the experiment as a control, to check that no crystals would nucleate in the supersaturated solution alone.

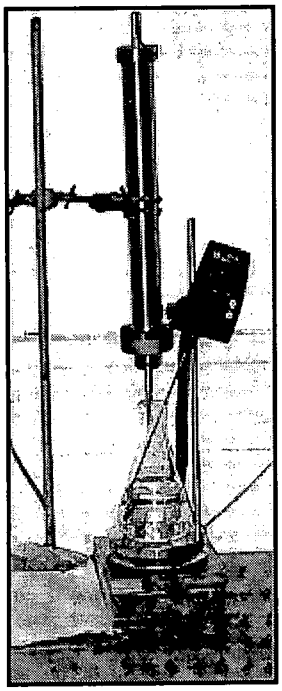


Figure 25 Filtering apparatus.

3.3.7 NUCLEATION EXPERIMENTS

Nucleation experiments were carried out by means of an automated Langmuir trough (NIMA Technology Ltd.), type 601M ($15 \times 7 \times 0.5$ cm^3), containing ≈ 70 cm^3 when filled so that a convex meniscus formed. The sapphire window present on the base of the trough made it ideal for the analysis *in situ* of crystal growth and shape. The apparatus was fitted with an optical microscope, which was in turn connected to a digital camera and to a TV screen. The formed crystals could so be visualised and photographed. The equipment is shown in Figure 26.

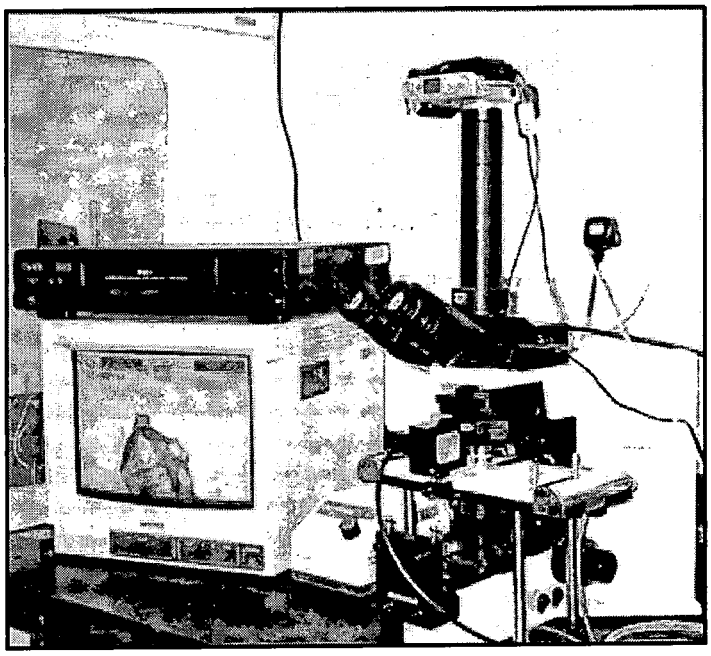


Figure 26 Langmuir trough (type 601M) fitted with the optical microscope connected to the digital camera and to the TV screen.

To avoid crystallisation upon the Wilhelmy plate, it was decide to use a thin platinum strip of set width. Before each experiment the plate was cleaned by dipping it into a concentrated hydrochloric acid solution and subsequent flaming.

7 μl of the film solution were added to the surface and, after evaporation of the chloroform, the film was compressed to the desired surface pressure. For all experiments it was decided to work with an initial surface pressure larger than zero. This was unavoidable, since any reduction in the amount of film material used would have required a compression of the film at too small areas in order to attain the desired Π. Nucleation under films at a small area would obviously have been disadvantageous. Moreover, the surface pressure was not maintained constant during the experiment because the continued movement of the barriers required to recompress the monolayer at the set Π could have disrupted the orientation of any

small crystal nuclei under the film and would have increased the loss of material from the surface.

The trough was covered throughout the 24 h of experiment with a plastic sheet to reduce the possibility of any impurity entering the system. Water evaporation from the trough during the experiment, which increased the dl-aspartic acid solution supersaturation and which was found to cause the loss of the solution up to 30%, was reduced by placing two 2 litres beakers of water kept at 70 °C under the sheet covering the trough.

3.4 CONCLUSIONS

Glycopolymers carrying peracetylated mono- and disaccharide residues, pAcGlcEMA, pAcGalMEA and pAcLacEMA, were found to form stable spreading Langmuir films on pure water at neutral and acidic pH. As expected for polymer films, Π-*A* curves were in all cases smooth. The glycopolymers bearing the monosaccharide residues gave monolayers characterised by limiting area per monomeric unit at high surface pressures approximately equal to 30 $Å^2$, whereas the same parameters was approximately double for the film formed by the polymer containing peracetylated lactose units in the side chains.

pAcGlcEMA and pAcGalEMA of approximately number-average molecular weight of 60,000 g mol^{-1} Langmuir films were found to act as efficient dl-aspartic acid nucleation promoters. In both cases crystals were obtained in high density and appeared to grow mainly bounded through the $\{\bar{1}11\}$ and $\{110\}$ faces and, in a smaller proportion, through the $\{\bar{2}02\}$ face. The faces grown attached to the films showed the numerous surface structure effects typical of epitaxial growth. They appeared enlarged compared to the corresponding faces of crystals grown

homogeneously from solution, as expected for the case of a diffusion limited growth. These faces also showed "hopper" growth, with the growing process proceeding outwards in all directions from an initial site of nucleation. pAcGalEMA with number-average molecular weight of approximately 30,000 g mol^{-1} also showed activity as a dl-aspartic acid nucleation catalyst, although with a lower efficiency compared to the analogue with higher molecular weight. The crystals nucleated beneath its Langmuir film appeared again to grow with faces $\{\bar{1}11\}$ and $\{110\}$ and, in a smaller proportion, $\{\bar{2}02\}$ attached to the monolayer. Crystals were less numerous but larger and more uniform in size relative to those grown beneath films of pAcGlcEMA and pAcGalEMA with $M_n \approx 60{,}000$ g mol^{-1}.

As expected from the parallelism between nucleation promotion activity of surfactants and crystal growth inhibition properties of the corresponding soluble analogues, the deprotected glycopolymers, pGlcEMA and pGalEMA, appeared to inhibit dl-aspartic acid growth. In particular the growth rate of faces $\{\bar{1}11\}$ and $\{110\}$ seemed to be reduced by the presence of the polymers in solution, resulting in crystals having morphologies dominated by these faces. Although only few crystal growth inhibition studies were carried out, the finding that the faces that grew attached to the monolayers of the protected polymers were the same as those whose growth was inhibited by the deprotected analogues suggests that pGlcEMA and pGalEMA may actually possess a certain crystal growth inhibition activity. Further crystal inhibition studies will enlighten this latter aspect.

Nucleation experiments carried out at different surface pressures, together with a detailed investigation of the structural properties of the films, could certainly provide information about the mechanism through which the monolayers induce the observed epitaxial growth. With regard to this, experiments where pAcLacEMA is

used as the film material (with the film possessing a limiting area per monomeric unit at high surface pressure double relative to pAcGlcEMA and pAcGalEMA films) would also give important additional data.

3.5 BIBLIOGRAPHICAL REFERENCES

[1] R. Popovitz-Biro, R. Edgar, I. Weissbuch, R. Lavie, S. Cohen, K. Kjaer, J. Als-Nielsen, E. Wassermann, L. Leiserowitz, M. Lahav, *Acta Polym.* **1998**, *49*, 626-635.

[2] S. W. J. Kuan, C. W. Frank, C. C. Fu, D. R. Allee, P. Maccagno, R. F. W. Pease, *J. Vac. Sci. Technol. B* **1988**, *6*, 2274-2279.

[3] S. W. J. Kuan, C. W. Frank, Y. H. Yen Lee, T. Eimori, D. R. Allee, R. F. W. Pease, R. Browning, *J. Vac. Sci. Technol. B* **1989**, *7*, 1745-1750.

[4] Y. K. Kim, M. H. Sohn, B. C. Sohn, E. Kim, S. D. Jung, *Thin Sol. Fi.* **1996**, *284-285*, 53-55.

[5] S. Mann, *Nature* **1988**, *332*, 119-124.

[6] R. J. Davey, S. J. Maginn, R. B. Steventon, J. M. Ellery, A. V. Murrell, J. Booth, A. D. Godwin, J. E. Rout, *Langmuir* **1994**, *10*, 1673-1675.

[7] I. Weissbuch, R. Popovitz-Biro, M. Lahav, L. Leiserowitz, *Acta Crystallogr.* **1995**, *B51*, 115-148.

[8] B. R. Heywood, S. Mann, *Adv. Mater.* **1994**, *6*, 9-20.

[9] S. J. Cooper, R. B. Session, S. D. Lubetkin, *Langmuir* **1997**, *13*, 7165-7172.

[10] R. Popovitz-Biro, J. L. Wang, J. Majewski, E. Shavit, L. Leiserowitz, M. Lahav, *J. Am. Chem. Soc.* **1994**, *116*, 1179-1191.

[11] E. M. Landau, M. Levanon, L. Leiserowitz, M. Lahav, J. Sagiv, *Nature* **1985**, *318*, 353-356.

[12] E. M. Landau, S. Grayer Wolf, M. Levanon, L. Leiserowitz, M. Lahav, J. Sagiv, *J. Am. Chem. Soc.* **1989**, *111*, 1436-1445.

[13] I. Weissbuch, G. Berkovic, R. Yam, J. Als-Nielsen, K. Kjaer, M. Lahav, L. Leiserowitz, *J. Phys. Chem.* **1995**, *99*, 6036-6045.

[14] R. Tang, Z. Tai, Y. Chao, *Chem. Lett.* **1996**, *7*, 535-536.

[15] B. R. Heywood, S. Mann, *Langmuir* **1992**, *8*, 1492-1498.

[16] F. Leveiller, D. Jaquemain, M. Lahav, L. Leiserowitz, M. Deutsch, K. Kjaer, J. Als-Nielsen, *Science* **1991**, *252*, 1532-1535.

[17] S. Mann, B. R. Heywood, S. Rajam, J. D. Birchall, *Nature* **1988**, *334*, 692-695.

[18] M. Gavish, R. Popovitz-Biro, M. Lahav, L. Leiserowitz, *Science* **1990**, *250*, 973-975.

[19] D. Jacquemain, F. Leveiller, S. P. Weinbach, M. Lahav, L. Leiserowitz, K. Kjaer, J. Als-Nielsen, *J. Am. Chem. Soc.* **1991**, *113*, 7684-7691.

[20] R. Popovitz-Biro, M. Lahav, L. Leiserowitz, *J. Am. Chem. Soc.* **1991**, *113*, 8943-8944.

[21] J. Majewski, R. Popovitz-Biro, W. G. Bouwman, K. Kjaer, J. Als-Nielsen, M. Lahav, L. Leiserowitz, *Chem. Eur. J.* **1995**, *1*, 304-311.

[22] D. J. Ahn, A. Berman, D. Charych, *J. Phys. Chem.* **1996**, *100*, 12455-12461.

[23] S. J. Cooper, R. B. Session, S. D. Lubetkin, *J. Am. Chem. Soc.* **1998**, *120*, 2090-2098.

[24] M. Lahav, I. Weissbuch, L. Addadi, L. Leiserowitz, *J. Am. Chem. Soc.* **1988**, *110*, 561-567.

[25] Y. Yeh, R. E. Feeney, *Chem. Rev.* **1996**, *96*, 601-617.

[26] M. M. Harding, L. G. Ward, A. D. J. Haymet, *Eur. J. Biochem.* **1999**, *264*, 653-665.

[27] S. Grandum, A. Yabe, K. Nakagomi, M. Tanaka, F. Takemura, Y. Kobayashi, P.-E. Frivik, *J. Cryst. Growth* **1999**, *205*, 382-390.

[28] T. Tsuda, S.-I. Nishimura, *J. Chem. Soc., Chem. Commun.* **1996**, 2779-2780.

[29] B. G. Davis, *Chem. Rev.* **2002**, *102*, 579-601.

[30] G. L. J. Gaines, *Insoluble Monolayers at Liquid-Gas Interfaces*, John Wiley & Sons, New York, **1966**, p. 187.

[31] G. L. J. Gaines, *Insoluble Monolayers at Liquid-Gas Interfaces*, John Wiley & Sons, New York, **1966**, pp. 264-274.

[32] L. Ter Minassian-Saraga, *J. Colloid Sci.* **1956**, *11*, 398-418.

[33] N. K. Adam, *The Physics and Chemistry of Surfaces*, Oxford University Press, London, **1941**.

[34] A. Roylance, T. G. Jones, *J. Appl. Chem.* **1961**, *11*, 329-334.

[35] W. F. Dunning, in *Chemistry of the Solid State* (Ed.: W. E. Garner), Butterworths Scientific Publications, London, **1955**, pp. 159-183.

[36] J. H. Brooks, A. E. Alexander, *Retardation of Evaporation of Monolayers*, Academic Press, New York, **1962**.

[37] I. Langmuir, *Science* **1936**, *84*, 379-383.

[38] I. Langmuir, V. J. Schaefer, *J. Am. Chem. Soc.* **1937**, *59*, 2400-2414.

[39] L. F. Chi, R. R. Johnston, H. Ringsdorf, *Langmuir* **1991**, *7*, 2323-2329.

[40] B. R. Pamplin, *Crystal Growth*, Pergamon, Oxford, **1975**.

[41] A. W. Adamson, *Physical Chemistry of Surfaces*, John Wiley & Sons, New York, **1982**.

4

PROPERTIES OF DEPROTECTED GLYCOPOLYMERS

4.1 INTRODUCTION

Protein-carbohydrate interactions are involved in a myriad of human biological processes, including the initiation of several kinds of infections and diseases, trafficking and clearance of glycoproteins, immune defence, fertilisation, recruitment of leukocytes to inflammatory sites, malignancy and metastasis[1-7]. Synthetic saccharides would therefore find wide employment in the investigation of biochemical phenomena proceeding through cellular recognition and in the design of new drug systems. Non-cytotoxic pharmaceuticals could be developed, consisting of molecules able to inhibit the initial pathogenic carbohydrate-mediated adhesion to the host cell. Alternatively, efficient macromolecular drug carriers designed to deliver the drug exclusively to those cell types which require the treatment, so minimising adverse side effects and reducing the cost of the therapy, could be produced. However, attempts to synthesise high-affinity specific compounds have been frustrated by the low intrinsic affinity of the protein-carbohydrate interaction. Typical binding constants of monovalent mono- and oligosaccharide ligands to specific lectins are in the range 10^3-10^6 M^{-1}[8]. Nature has overcome the intrinsic weakness of binding through multivalency: lectins typically exist as oligomeric structures in physiological conditions. Therefore, the preparation of multivalent compounds able to bind to multivalent targets seems a promising strategy in the search for high-affinity protein ligands. The activity enhancement of a multivalent ligand compared to the

corresponding monovalent saccharide on a valency-corrected basis, defined as the "cluster glycoside effect"[9-11], is nowadays generally accepted. However, the physical origin of the phenomenon is still not well understood[12]. It has been repeatedly pointed out that both structural and energetic aspects of the lectin-carbohydrate interaction have to be elucidated, if a deep understanding of the process is to be achieved[7,13,14]. From an energetic point of view, the delicate balance between enthalpy and entropy needs to be revealed. Most techniques currently in use for the determination of the thermodynamic parameters of protein-ligand association directly provide only K values, which give little information about the forces involved in the complexation process. ΔH and ΔS are derived only indirectly by means of the van't Hoff equation and the assumptions required for the use of this equation question the accuracy of the values obtained. Isothermal titration microcalorimetry, ITM, is the only technique which allows the determination of the thermodynamic binding parameters K, ΔG, ΔH, ΔS and n (stoichiometry of binding, carbohydrate:protein) in a single experiment.

Few calorimetric studies of multivalent carbohydrate-protein interactions[11,12,15-21] (indeed few for any multivalent ligand[22,23]) have been reported so far. Moreover, ITM has been limited to small or dendritic glycoconjugates possessing a maximum number of carbohydrate residues equal to six. Despite hemagglutination inhibition studies (HIA), enzyme-linked lectin assays (ELLA), enzyme-linked immunosorbent assays (ELISA) and surface plasmon resonance (SPR) having shown the great potential of glycopolymers as high-affinity polyvalent ligands, calorimetrically derived thermodynamic data of their association with lectins have never been reported. The reasons may reside in the complexity of the system and the difficulties associated with the exact knowledge of the ligand composition and

structural features. Thus, high pure and strictly well-defined polymers are absolutely needed for a true understanding of the results obtained. Furthermore, polyvalent ligand-lectin interactions often proceed through an aggregative mechanism in which each ligand cross-links distinct protein molecules. This phenomenon may result in the precipitation of the aggregates, which, in turn, may hamper the ITM experiment. Thermodynamic parameters are, in fact, state functions and the enthalpy recorded is the sum of all processes occurring in the reaction cell.

In this study, for the first time the interaction of polymeric glycoconjugates with a lectin was investigated by isothermal titration microcalorimetry. In particular, the binding properties of the glycopolymer carrying the galactose unit to peanut agglutinin (PNA, lectin from *Arachis Hypogaea*) were determined. As discussed in Chapter 2, deprotected polymers were obtained following two routes (designated A and B): polymerisation of the deprotected glycomonomers (route A) and deprotection of the corresponding polymers (route B). A thorough investigation of the materials' properties revealed that the procedure of preparation strongly affected the polymer composition. Whilst the deprotection of the polymers (route B) resulted in an incomplete deacetylation and yielded products of ill-defined compostion, fully deacetylated, well-defined and pure materials could be obtained by polymerisation of the corresponding deprotected glycomonomers (route A). It was therefore interesting to test whether the polymer composition affected its binding properties.

Prior to proceeding with the evaluation of the thermodynamic binding parameters, the behaviour of the deprotected polymers in aqueous solutions was investigated by dynamic light scattering (DLS). pGlcEMA-A,B and pGalEMA-A,B were found to self-assemble in water and buffer solutions. The aggregates formed were analysed by transmission electron microscopy (TEM) and environmental

scanning electron microscopy (ESEM). Considering that the enthalpy determined by ITM is the sum of all processes which take place in the sample cell, it is fundamental to assess all the phenomena that can contribute to the recorded heat value.

The carbohydrate-protein interaction was probed in two ways. Initially, qualitative information was obtained by ultraviolet-difference spectroscopy. Due to the presence of tyrosine residues in the binding site[24], PNA develops a strong ultraviolet-difference spectrum on ligand binding[25-28]. The presence of UV-difference spectra indicated that all ligands bound to the lectin.

Subsequently, calorimetric quantitative determinations of the thermodynamics of binding of D-galactose, GalEMA and pGalEMA-A,B for peanut agglutinin were performed.

4.1.1 General Aspects of Isothermal Titration Microcalorimetry, ITM

Isothermal titration microcalorimetry (ITM) allows the measurement of the heat flow in chemical, physical and biological processes and, therefore, their relative thermodynamic binding parameters[29].

ITM has been applied to a wide range of areas, including biological research and chemical and pharmaceutical industry[30], in order to determine the binding affinity of numerous types of interactions, such as nucleic acid-protein[31], protein-protein[32-34], antigen-antibody[35-37] and protein-carbohydrate[11,12,15-20,38-47].

A schematic representation of a titration microcalorimeter is reported in Figure 1.

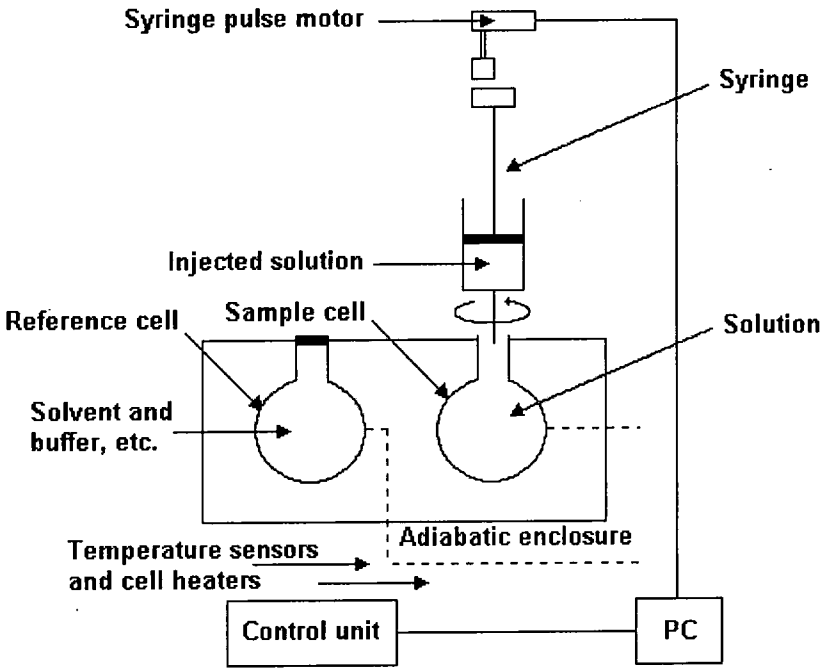


Figure 1 Titration microcalorimeter showing reference cell and sample cell; the latter is fitted with a stirred syringe for injecting aliquots of a solution into the solution held in the sample cell.

The sample and reference cells are brought to thermal equilibrium and then heated at a constant rate. Addition of the ligand to the sample cell perturbs this equilibrium in such a way depending on the enthalpy of the interaction. The resulting difference in temperature is detected by highly sensitive thermopiles (usually capable, as in the present experiment, to respond to temperature gradients of less than one millionth of a degree Celsius). As a result, a voltage proportional to the rate of heat energy transfer across the thermopiles is generated and the reference and sample cells are brought back to equilibrium. The reference cell is placed in parallel to the sample cell and it is also connected to the thermopiles. It contains the same volume as the sample cell. Therefore, any disturbances from the surroundings will affect both sample and reference cells to the same extent and they will be subtracted from the total heat output.

Typically, for lectin-carbohydrate interactions, the lectin is placed in the sample cell and the ligand is added in a series of injections.

Raw data are in units of power *versus* time. Integration with respect to time provides the binding isotherm, relating the heat to the ligand concentration. Assuming a model, the enthalpy (ΔH), the binding constant (K) and the stoichiometry (n) of the process are determined by non-linear least-squares analysis of the binding isotherm. The free energy of complex formation (ΔG) is then calculated from $\Delta G = -RT \ln K$ and the entropy of binding determined as $T\Delta S = \Delta H - \Delta G$.

4.1.2 PEANUT AGGLUTININ, LECTIN FROM *ARACHIS HYPOGAEA*

Peanut agglutinin is a homotetrameric, non-glycosylated leguminous lectin. It is terminal D-galactosyl residue specific, with a particularly high affinity for the tumour-associated T-antigenic disaccharide [DGalβ(1→3)DGalNAc, Thomsen-Friedenreich antigen]. Each subunit, M_r 27,000 ± 1,500 Da[48], contains 236 amino acid residues and presents the well-known jelly roll tertiary structural fold, characteristic of all legume lectins[49] (Figure 2b). The tertiary structure is constituted by a six-stranded nearly flat sheet (sheet 1), a curved seven-stranded sheet (sheet 2), a small five-stranded sheet (sheet 3) which plays an important role in connecting the two larger sheets, and a number of loops which make up 54% of the polypeptide chain. The arrangement of the loops gives rise to a second hydrophobic core between sheet 2 and the loops in addition to that between sheets 1 and 2[50] (Figure 3). Differences between the tertiary structures of peanut agglutinin and other legume lectins are largely confined to loops and chain terminii[51]. The two metal ions, Ca^{2+} and Mn^{2+}, are located at the same positions as in other legume lectins.

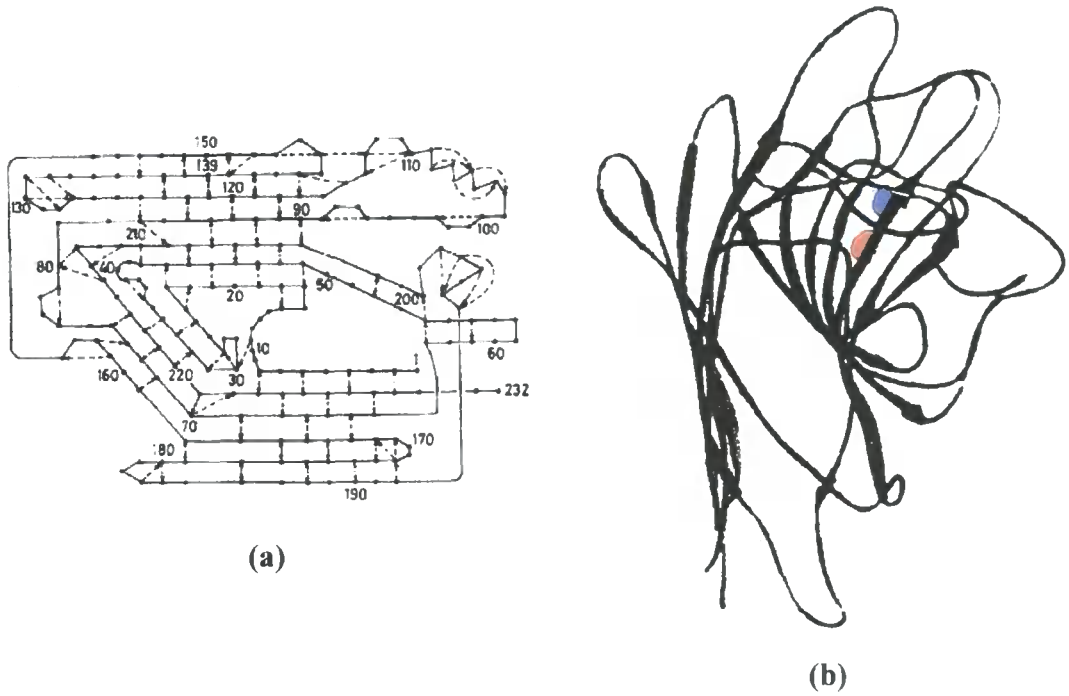


Figure 2 a), Schematic representation of the main-chain to main-chain hydrogen bonds present in all subunits. A hydrogen bond which connects Gly103 N and Gly210 O is not shown in the Figure; b), ribbon drawing of a subunit in peanut lectin. The red and blue circles represent manganese and calcium ions, respectively.

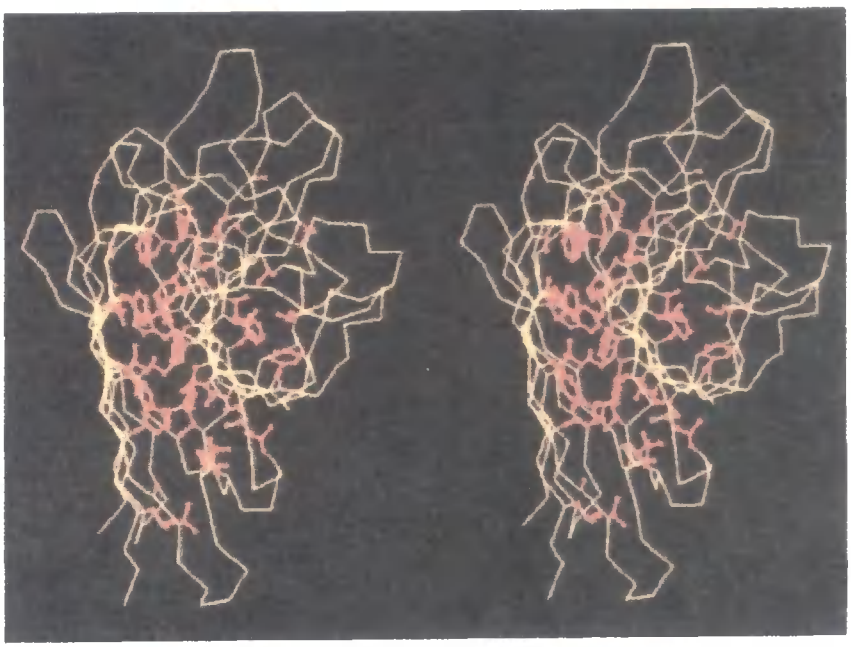


Figure 3 Stereo view of the buried hydrophobic residues (red) in the subunit with the main-chain shown in yellow.

Despite having the same tertiary structure at the monomeric level, legume lectins exhibit considerable variation in their quaternary arrangement arising from

small changes in their sequence. Different modes of dimerisation are possible, resulting in the formation of the canonical dimer* observed in concanavalin A, pea lectin, favin and *L. ochrus* or of the GS4 dimer characteristic of *G. simplicifolia* and *E. corallodendron*[52]. Further association of the dimers into tetramers is also observed, for instance in concanavalin A and peanut agglutinin. Unlike all the other well-characterised tetrameric proteins constituted by identical subunits, PNA has neither 222 (D_2) nor 4-fold (C_4) symmetry. Instead, it exhibits the unusual "open" assembly reported in Figure 4[51].

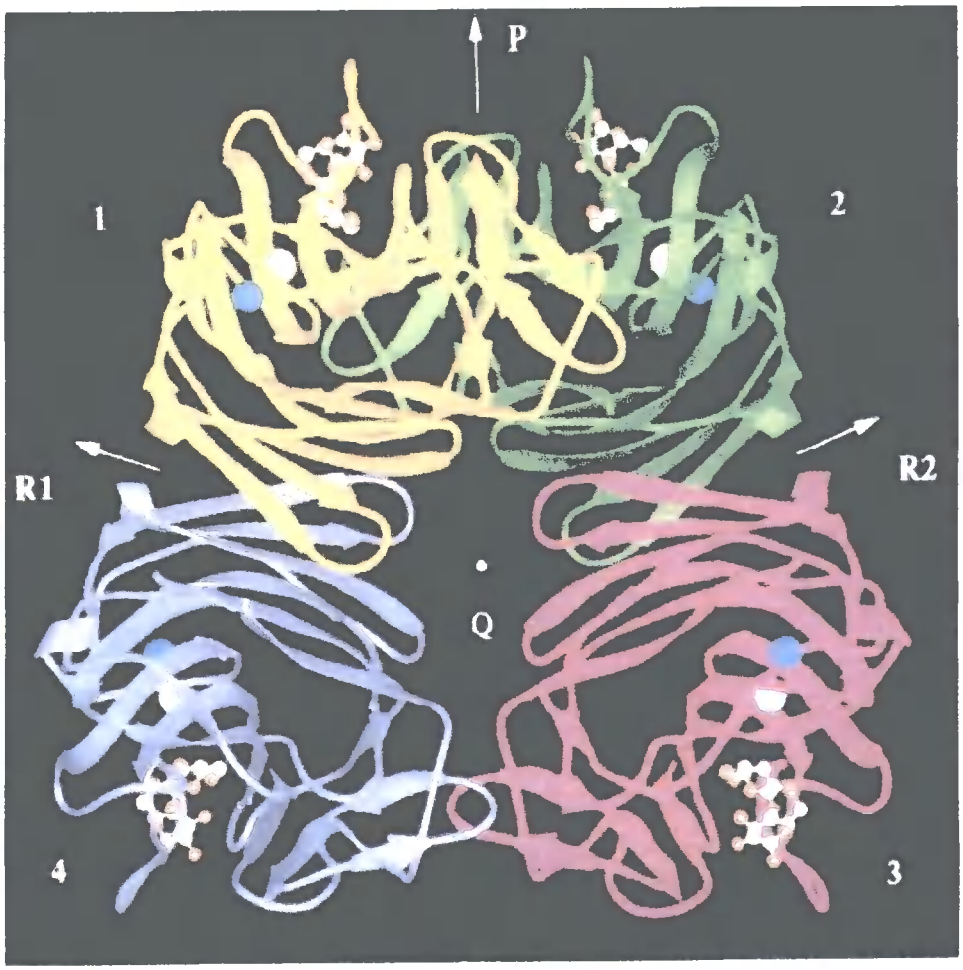


Figure 4 Quaternary structure of PNA complexed with lactose. Blue and white circles represent manganese and calcium ions, respectively.

* See Chapter 1, Section 1.1.1.1.

Despite the lack of D_2 symmetry, it has been shown that peanut lectin tetramer is a dimer of a dimer[53], where the two dimers, 1-4 and 2-3, are constituted from two GS4 arrangements. The open quaternary association is stabilised mainly by hydrophobic, hydrogen-bonded and water-mediated interactions. While the dimerisation process results in the burial of 1920 $Å^2$ surface area of which 71% is non-polar, the percentage of hydrophobic surface area buried during the further formation of the tetramer is relatively low. The dimers interact in such a way that monomers 1 and 2 associate in a canonical fashion except that the sheets do not interact directly, but by means of six water bridges. As a result, the 1-2 interface is intrinsically less stable than the analogous interface in Con A[52]. As with most plant lectins, the quaternary structure of PNA depends on the pH. The lectin is a tetramer at physiological pH[25], which dissociates reversibly into dimers at pH below 5.1. Below pH 3.4 the lectin is totally dimeric[54].

The protein in association with lactose was first crystallised at neutral pH in an orthorhombic form[55] and then at acidic pH in two monoclinic forms and one triclinic form[56]. Subsequently, crystal structures of the complexes of the protein with methyl-β-galactose (MeGal)[57], the T-antigenic disaccharide (Tant)[58,59], *N*-acetyllactosamine (LacNAc)[57] and lactose (Lac)[60] were reported (Figure 5). Residues in four loops – 75 to 83, 91 to 106, 125 to 135 and 211 to 216 – make up the carbohydrate binding pocket toward the top right of the subunits in legume lectins[49]. A triad of amino acid residues is involved in the interaction with ligands in almost all legume lectins: Asp83, Gly104 (present in all legume lectins with the exception of Con A) and Asn127 (in the PNA numbering scheme). Another invariant feature of the sugar-binding sites is the presence of an aromatic residue[61] corresponding to Tyr125 in PNA, which stacks against the sugar ring. These invariant features belong to the first

three loops mentioned, while interactions involving amino acid residues in loop 211-216 appear to be primarily responsible for the sugar specificity of each legume lectin. In fact, this loop is highly variable in terms of length, sequence and conformation.

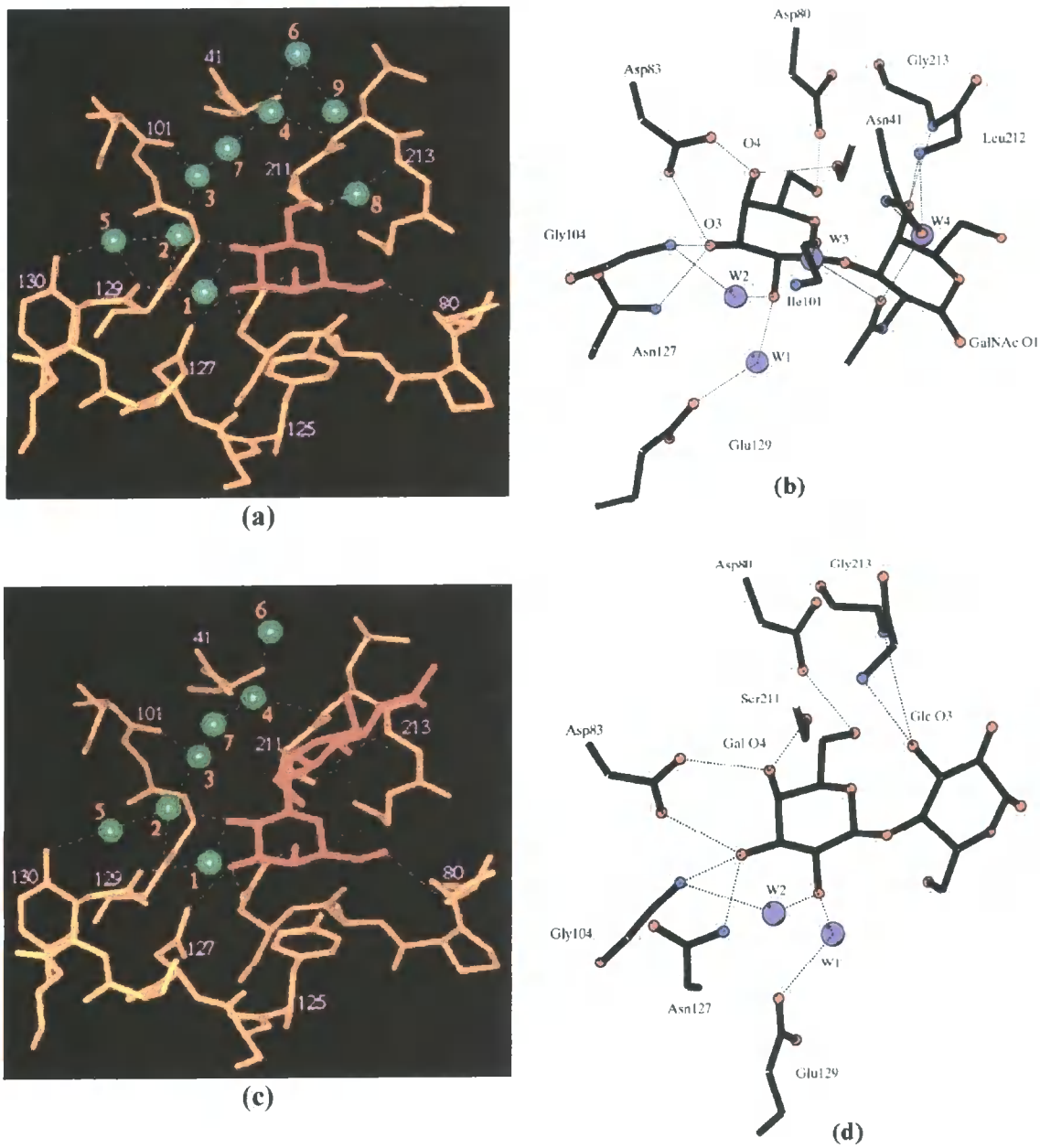


Figure 5 a), Stereo diagram of PNA-MeGal complex. The sugar is represented in red. The interacting residues (yellow-orange) are numbered in pink. Hydrogen bonds are represented by pink dotted lines. The water molecules (green spheres) are numbered in orange; b), crystal structure of PNA-Tant complex. Water molecules are represented by violet spehers; c), stereo diagram of PNA-LacNAc complex. Colours as for 5a; d), crystal structure of PNA-Lac complex. Colours as for 5b.

The four invariant hydrogen bonds Asp83 OD1-Gal O3, Asp83 OD2-Gal O4, Gly104 N-Gal O3 and Asn127 ND2-Gal O3 can be seen in Figures 5a-d. The stacking interaction with the aromatic side chain of Tyr125 is also shown in the stereo diagrams of PNA-MeGal and PNA-LacNAc complexes. The hydrogen bond between Asp80 OD2 and Gal O6 can also be observed in all four complexes. Further hydrogen bonds involve Ser211 OG, which belongs to the fourth loop (211-216). Water bridges involving W1 and W2 occur in all four complexes: Gal O2 is connected to Glu129 OE2 through W1 and to Gly104 N through W2. Gal O1 is connected to Ser211 OG and to Gly213 N through W8 in PNA-MeGal. The residues within 4 Å from the sugars in PNA-MeGal, PNA-LacNAc and PNA-Lac complexes are listed in Table 1, together with the formed hydrogen bonds. Moreover, about 40 van der Waals contacts between these residues and the galactose moiety are made in each complex. As a result, the position of the galactose ring is approximately the same in all cases examined. The number of residues within 4 Å from the carbohydrate increases by one when the second sugar is added, with the number of van der Waals contacts raising to 60. PNA binds the T-antigen 20 times more strongly than it binds lactose. As already said, the interactions between the galactose moiety and the protein are similar in all four complexes. As can be seen from Figure 5, the conformation of the disaccharides is such that O4 of the second ring of the T-antigen occupies the same position as O3 of the glucose moiety in PNA-Lac. O4 is therefore hydrogen bonded to Ser211 OG and Gly213 N instead of O3 as in the other two disaccharides. The major additional interactions between the T-antigen and PNA, compared to the other disaccharides, occur through two water molecules involving the carbonyl O atom of the acetamido group: W3 connects the O atom to Ile101 O, while W4 connects it to Asn41 ND2 and to Leu212 N. These two water molecules also exist in the other complexes, but in

these cases they interact only with the protein, due to the lack of sugar atoms in their immediate vicinity. Therefore, the high specificity of PNA for the T-antigenic disaccharide appears to be generated by two water bridges, which illustrates the critical importance of water-mediated interactions in protein-carbohydrate binding.

Table 1 Protein-sugars interactions.

Protein atom	Sugar atom	Distance (Å) in subunit			
		1	2	3	4
Metyl-β-galactose					
Hydrogen bonds					
Asp83 OD1	Gal O3	2.6	2.7	2.7	2.6
Gly104 N	Gal O3	3.2	3.0	2.9	2.9
Asn127 ND2	Gal O3	2.8	2.9	2.9	2.8
Asp83 OD2	Gal O4	2.7	2.7	2.8	2.7
Ser211 OG	Gal O4	2.8	2.7	2.8	2.7
Ser211 OG	Gal O5	3.2	3.5	3.6	3.3
Asp80 OD2	Gal O6	3.0	3.2	3.3	3.0
Residues within 4 Å from the sugar					
Asp80, Ala82, Asp83, Gly103, Gly104, Tyr125, Asn127, Ser211 and Gly213.					
***N*-acetyllactosamine**					
Hydrogen bonds					
Asp83 OD1	Gal O3	2.6	2.5	2.8	2.8
Gly104 N	Gal O3	3.2	3.4	3.0	2.0
Asn127 ND2	Gal O3	2.9	2.8	2.8	2.7
Asp83 OD2	Gal O4	2.7	2.8	3.0	2.8
Ser211 OG	Gal O4	2.7	2.8	2.9	3.0
Ser211 OG	Gal O5	3.0	3.3	3.6	3.3
Asp80 OD2	Gal O6	3.2	3.5	3.3	3.4
Ser211 OG	GlcNAc O3	3.1	3.3	3.6	3.2
Gly213 N	GlcNAc O3	2.7	2.9	3.3	2.7
Residues within 4 Å from the sugar					
Asp80, Ala82, Asp83, Gly103, Gly104, Tyr125, Asn127, Ser211, Leu212 and Gly213.					
Lactose					
Hydrogen bonds					
Asp83 OD1	Gal O3	2.67	2.43	2.58	2.49
Gly104 N	Gal O3	3.08	2.92	3.07	2.81
Asn127 ND2	Gal O3	2.86	3.00	3.12	3.02
Asp83 OD2	Gal O4	2.59	2.68	2.64	2.55
Ser211 OG	Gal O4	2.62	2.82	2.66	2.76
Ser211 OG	Gal O5	3.12	3.34	3.09	3.16
Asp80 OD2	Gal O6	3.33	3.39	3.36	2.98
Ser211 OG	Glc O3	3.34	3.31	3.58	2.98
Gly 213 N	Glc O3	2.92	2.98	3.29	3.28
Residues within 4 Å from the sugar					
Asp80, Ala82, Asp83, Gly103, Gly104, Tyr125, Asn127, Ser211, Leu212 and Gly213.					

A detailed study of the water molecules in the combining sites of the four complexes has been carried out[57]. The 16 subunits of the complexes, along with the hydration shells, were superimposed. All the water molecules within 5 Å from any carbohydrate atom were considered to belong to the binding site. Such superimposition is reported in Figure 6.

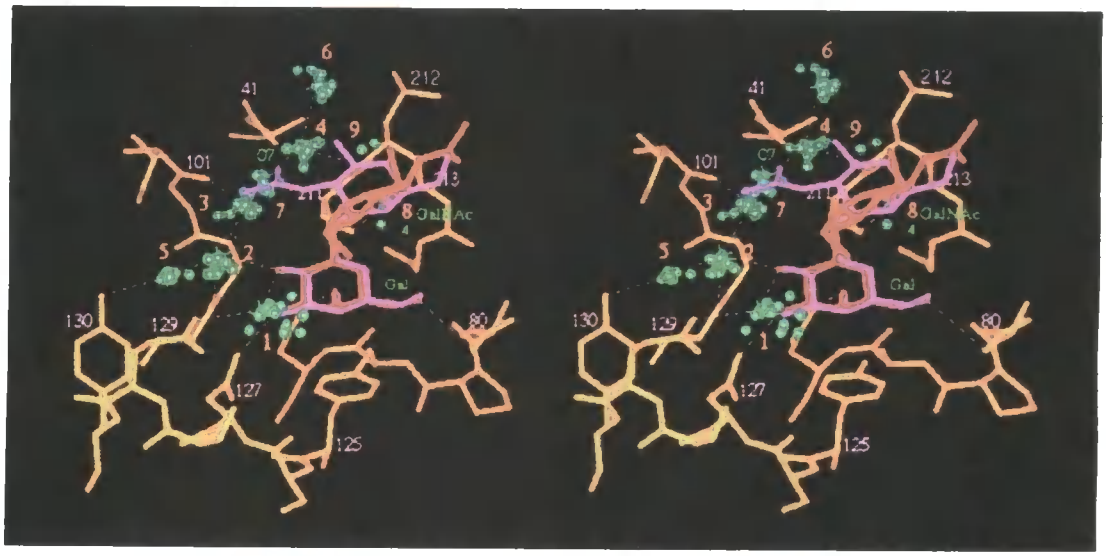


Figure 6 Stereo diagrams of the water clusters (green), numbered in orange, in the binding site after superimposition of the 16 subunits in the four complexes of PNA. The numbering in green is for the T-antigenic disaccharide (pink). The *N*-acetyllactosamine is coloured in red. The surrounding peptide stretches (yellow-orange) are numbered in pink. Hydrogen bonds are represented by pink dotted lines.

Most water molecules appeared in clusters, indicating their relative invariance in the binding sites. A cluster was defined as a set of water molecules from different subunits in which all water O atoms are within a sphere of 1.5 Å radius with the average position as the centre. Water clusters W1-W6 are not displaced by sugar molecules in any of the four complexes. The acetamido group displaces W7 in the T-antigen complex. W8 and W9 are displaced by the second ring and they occur only in PNA-MeGal. W1, W2 and W5 form a network which bridges galactose and PNA at the primary binding site. W3 and W4 are also hydrogen bonded to the protein: W3 to Ile101 O and W4 to Asn41 ND2 and Leu 212 N. In PNA-Tant both hydrogen bond to

GalNAc O7, giving rise to the two bridges responsible for the high specificity of the lectin for this disaccharide. W6 is hydrogen bonded to W4, but not to any protein or sugar atom. Interestingly, as already reported for other carbohydrate-binding proteins[62], ordered water molecules can be found in the unligated form at positions corresponding to hydroxyl groups in the ligated form (for instance, a water molecule occupies the postion of GalNAc O7 of the T-antigen). In general, water molecules in the carbohydrate-binding region mimic the ligand to a substantial extent not only at the primary site, but also in the regions adjacent to it. Molecular dynamics simulations carried out for PNA-Tant and PNA-Lac complexes showed that the number of alternative binding modes is higher for the former than for the latter, resulting in the "breaking" of enthalpy-entropy compensation characteristic of the binding of monovalent sugars to lectins[60]. Moreover, mutational analysis at Leu212[63] and Asn41[64] proved the critical role of these two amino acid residues in the binding of the T-antigenic disaccharide, strengthening the theory that non primary contacts, generally made through water molecules, are involved in imparting extraordinary specificity.

The interaction of mono- and disaccharides with peanut agglutinin has been investigated by hemagglutination inhibition studies, fluorimetry, ultraviolet-difference spectroscopy, C-13 NMR, circular dichroism (CD) and molecular dynamics calculations (MD)[25-28,47,48,60,65-68]. The results reported from different groups showed good agreement. As already mentioned, the highest affinity was determined for the tumour-associated T-antigenic disaccharide, while the higher affinity for methylgalactopyranosides compared to galactose was attributed to the binding of sugars in the pyranose form[48]. Between the galactosides, a slight preference for the α-anomers was detected, as the α-linkage is more sterically acceptable[69]. The

affinity increased with decreasing the temperature, especially for disaccharides. PNA in dimeric form (low pH) binds sugars with the same stoichiometry as the tetramer (one binding site per protomer) but with an association constant one order of magnitude lower than that for the tetramer[25,54]. Thermodynamic parameters determined by means of UV-difference spectroscopy, fluorimetry and isothermal titration microcalorimetry are reported in Table 2. In all cases, the interaction is enthalpically driven, as a result of the formation of direct and water-mediated hydrogen bonds and van der Waals interactions, and shows the typical enthalpy-entropy compensation[70]†.

Table 2 Thermodynamic binding parameters of PNA-monovalent ligand interactions.

Ligand	$10^{-3} \times K$ M^{-1}	ΔG kJ mol^{-1}	ΔH kJ mol^{-1}	ΔS J mol^{-1} K^{-1}
Methyl-α-D-Gal	1.8 [a]	-17.6 [a]		
	1.8 ± 0.1 [b]	-18.5 ± 0.1 [b]	-42 ± 2 [b]	-79 ± 6 [b]
	3.084 ± 0.127 [d]			
Methyl-β-D-Gal	1.4 [a]	-17.2 [a]		
	1.0 ± 0.1 [b]	-17.1 ± 0.1 [b]	-43 ± 3 [b]	-87 ± 9 [b]
	1.870 ± 0.078 [d]			
D-Galactose	1.2 [a]	-16.7 [a]		
Lactose	1.5 [a]	-17.2 [a]		
	1.3 ± 0.1 [b]			
	2.0 [e]	-18.8 [e]	-40.6 [e]	-73 [e]
Methyl-β-D-Lac	1.4 ± 0.1 [b]	-17.9 ± 0.1 [b]	-65 ± 4 [b]	-156 ± 14 [b]
β-D-Gal(1→3)-D-GalNAc	28 ± 2 [b]	-25.4 ± 0.2 [b]	-78 ± 5 [b]	-177 ± 16 [b]
T-antigen	20.6 [c]	-24.7 [c]	-59.0 [c]	-115 [c]
MeUmb-α-D-Gal	1.7 ± 0.3 [c]	-18.4 [c]		
MeUmb-β-D-Gal(1→3)-D-GalNAc	33 ± 4 [c]	-25.5 [c]		

[a] Phosphate buffer, containing 200 mM NaCl, 10 °C, determined by UV-difference spectroscopy[28]. [b] PBS, pH 7.2, 25 °C (for Methyl-α-D-Gal, Methyl-β-D-Gal and Lactose)[26]; 10 mM phosphate buffer, containing 150 mM NaCl, pH 7.2, 25 °C [for Methyl-β-D-Lac and β-D-Gal(1→3)-D-GalNAc], determined by UV-difference spectroscopy[27]. [c] 50 mM Tris-HCl buffer, containing 500 mM $MgCl_2$, 1 mM $CaCl_2$, 1 mM $MnCl_2$, pH 6.9, determined by fluorimetry[67]. [d] Phosphate buffer, pH 7.4, 25 °C, determined by ITM[47]. [e] PBS, pH 7.4, 297.0 K for lactose and 297.9 K for T-antigen, determined by ITM[60].

† See Chapter 1, Section 1.3.2.1.

4.2 RESULTS AND DISCUSSION

4.2.1 SELF-ASSOCIATION OF DEPROTECTED GLYCOPOLYMERS IN AQUEOUS MEDIA

pGlcEMA-A,B (**5a-A,B**) and pGalEMA-A,B (**5b-A,B**) tended to form aggregates in aqueous media. The aggregation of polymers was investigated by dynamic light scattering (DLS) and the critical aggregation concentration (CAC) determined. Curves of scattering intensity *versus* polymer concentration are reported in Figure 7.

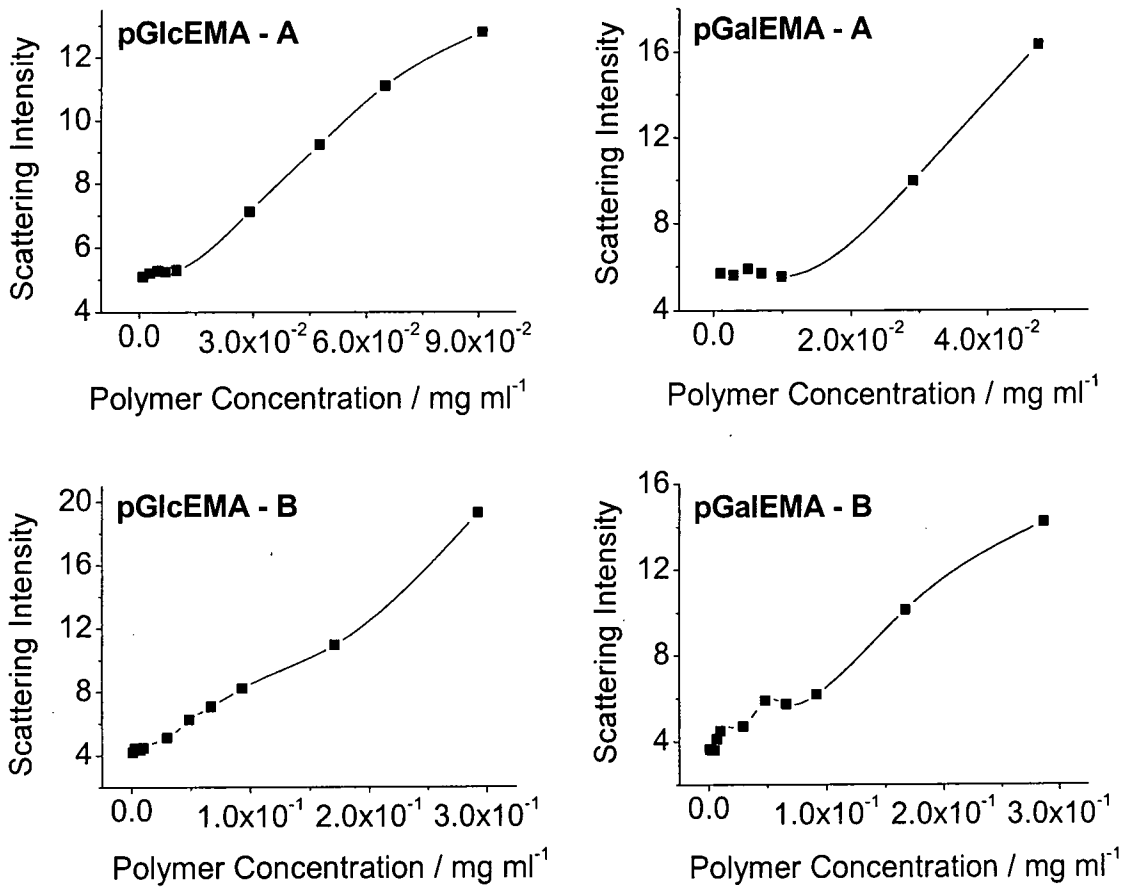


Figure 7 Scattering intensity *versus* polymer concentration as determined by DLS.

It can be noted that the curves for deprotected polymers obtained by route A (polymerisation of the corresponding deprotected monomers) exhibited the typical behaviour of surfactants aggregating in aqueous media with a starting plateau and an increase of the scattering intensity corresponding to the critical aggregation concentration. CACs of pGlcEMA-A and pGalEMA-A were both *ca.* 1×10^{-2} mg ml^{-1}. On the contrary, the behaviour found for pGlcEMA-B and pGalEMA-B could be explained considering the presence of partially protected polymers that did not form mixed aggregates with the fully deprotected ones. As previously reported the deprotection of the polymers (route B) was not quantitative.

The morphology and the size distribution of the aggregates were investigated by transmission electron microscopy, TEM (Figure 8) and DLS (Figure 9), respectively. In all cases, the aggregates showed a spherical morphology and a core-shell architecture, with size varying on a broad range – 15-500 nm –, as a result of the polydispersity of the polymers (Table 3). It can be noted that in general polymers with the highest molecular weights formed the largest aggregates.

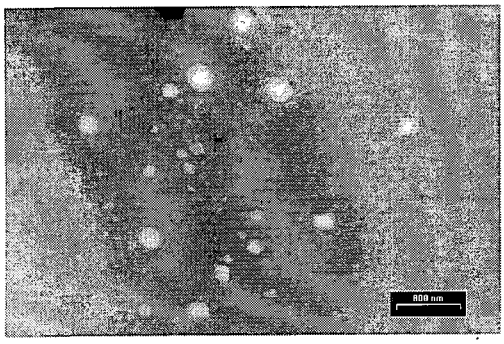
Table 3 Number- and weight-average molecular weights, dn/dc and effective aggregate diameter of deprotected polymers[‡].

Polymer	M_n g mol^{-1}	M_w g mol^{-1}	M_w M_n	dn/dc ml g^{-1}	d (aggregates) nm
pGlcEMA-A	23,000	60,000	2.61	0.131	21.5
pGalEMA-A	461,000	1,019,000	2.21	0.140	42.4
pGlcEMA-B	5,932	10,410	1.75	0.126	20.7
pGalEMA-B	2,984	6,416	2.15	0.146	9.4

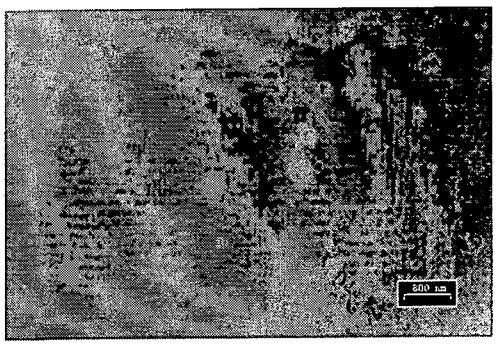
The larger aggregates size observed by TEM compared to DLS was expected, considering that TEM measurements were carried out using solutions 10 times more concentrated than those employed for DLS. Nevertheless, the size and morphology

[‡] For the details of molecular weight measurements, see Chapter 2.

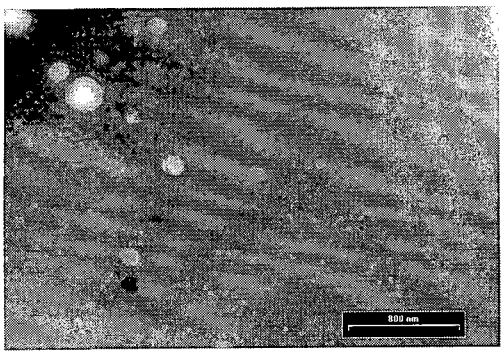
determined by TEM could be affected by the drying process required for the preparation of the samples: flattening of the particles could alter their real features.



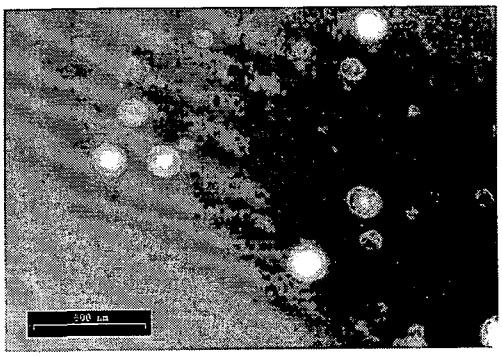
(a) pGlcEMA-A



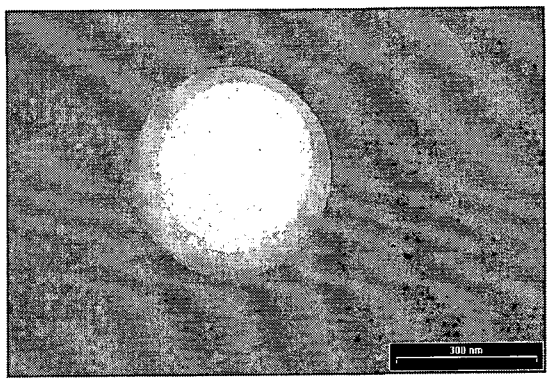
(b) pGalEMA-A



(c) pGlcEMA-B



(d) pGalEMA-B



(e) CORE-SHELL ARCHITECTURE

Figure 8 a-d), TEM micrographs obtained using 10 mg ml^{-1} aqueous solutions (bars 800 nm); e), TEM micrograph showing the core-shell structure observed for all aggregates (bar 300 nm).

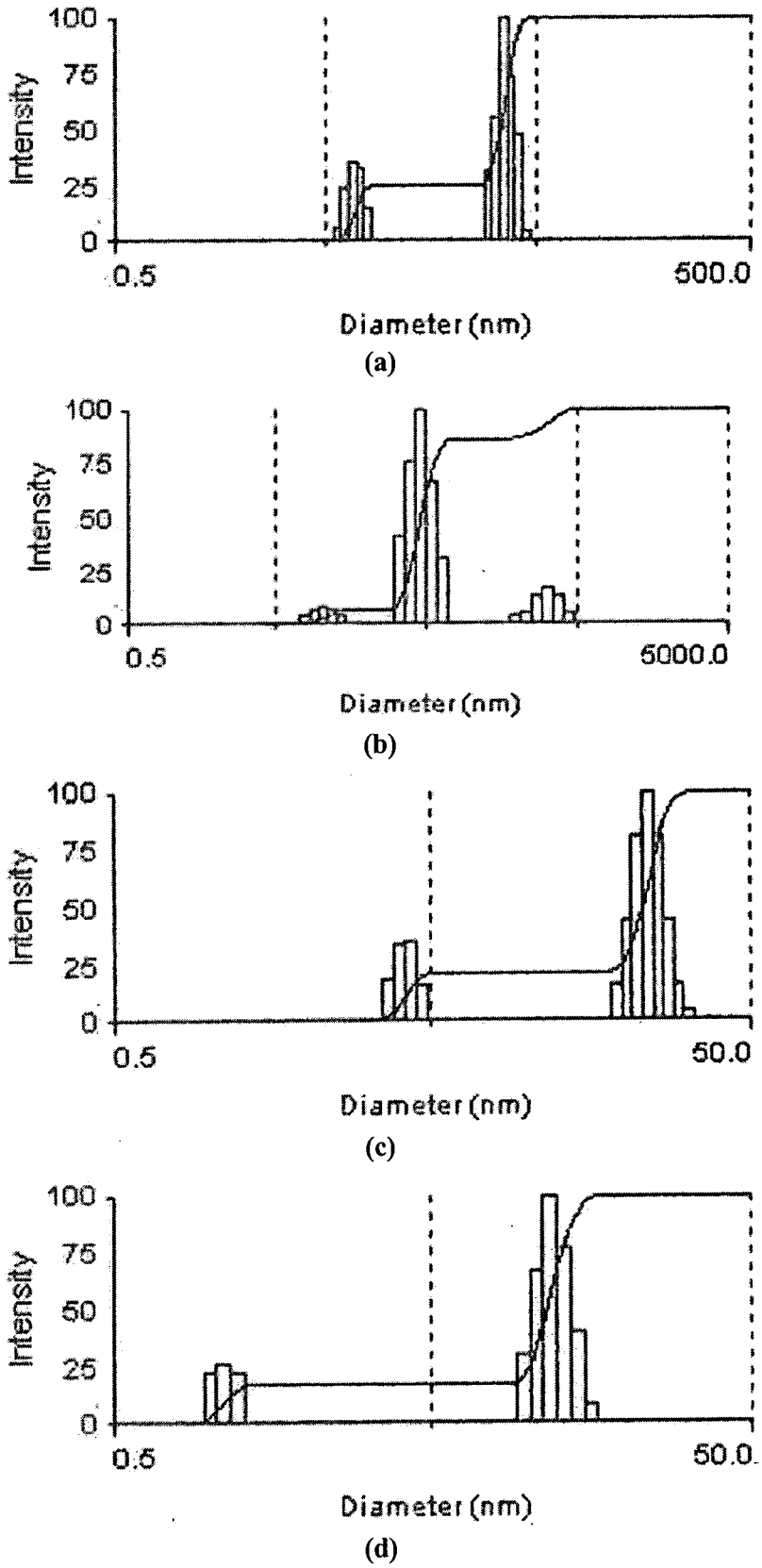


Figure 9 Particle size distribution as determined by DLS, polymer concentration equal to 1 mg ml^{-1} in water at 25 °C. a), pGlcEMA-A; b), pGalEMA-A; c), pGlcEMA-B; d), pGalEMA-B.

The observed core-shell architecture is hardly understandable considering the structure of the polymers. Moreover, fluorescence studies carried out using pyrene as hydrophobic probe showed no variations in the emission spectrum of pyrene in the presence of the polymers. In principle, the emission spectrum of pyrene changes depending on the polarity of the environment. For surfactants forming micelle-like assemblies, on formation of aggregates pyrene is encapsulated into the hydrophobic core of the aggregates, with subsequent increase of the intensity of band III (λ_{III} = 384 nm) and decrease of the band I/band III ratio (λ_{I} = 373 nm). Pyrene is largely used for the determination of the critical micelle concentration (CMC) of surfactants. The aggregation of pGlcEMA and its CAC in aqueous solution has been already determined by fluorimetry[71]. However, measurements carried out here did not show any variation of the pyrene emission spectrum, indicating that the probe environment did not change when polymer aggregates were present in solution (Figure 10).

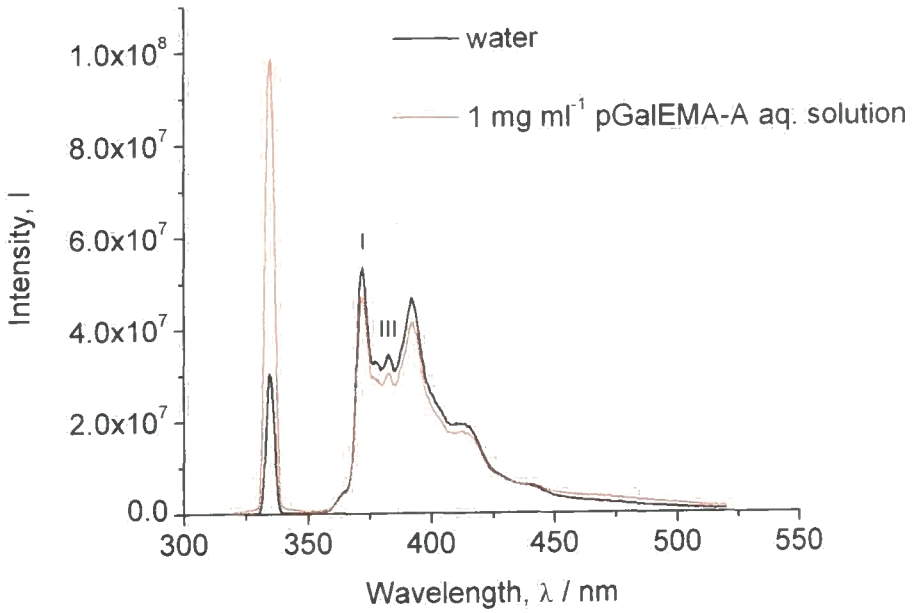


Figure 10 Emission spectra of pyrene in water and in pGalEMA-A aqueous solution. λ_{excit} = 334 nm.

Measurements were performed following several procedures for the sample preparation, including that reported for pGlcEMA, but results did not show any entrapment of pyrene into hydrophobic regions. Therefore, pGlcEMA and pGalEMA particles might be better envisaged as loose aggregates, as a result of inter- and intramolecular direct and water-mediated hydrogen bonds. Such aggregates have been reported and extensively studied for polyethylene oxide, PEO[72] (Figure 11).

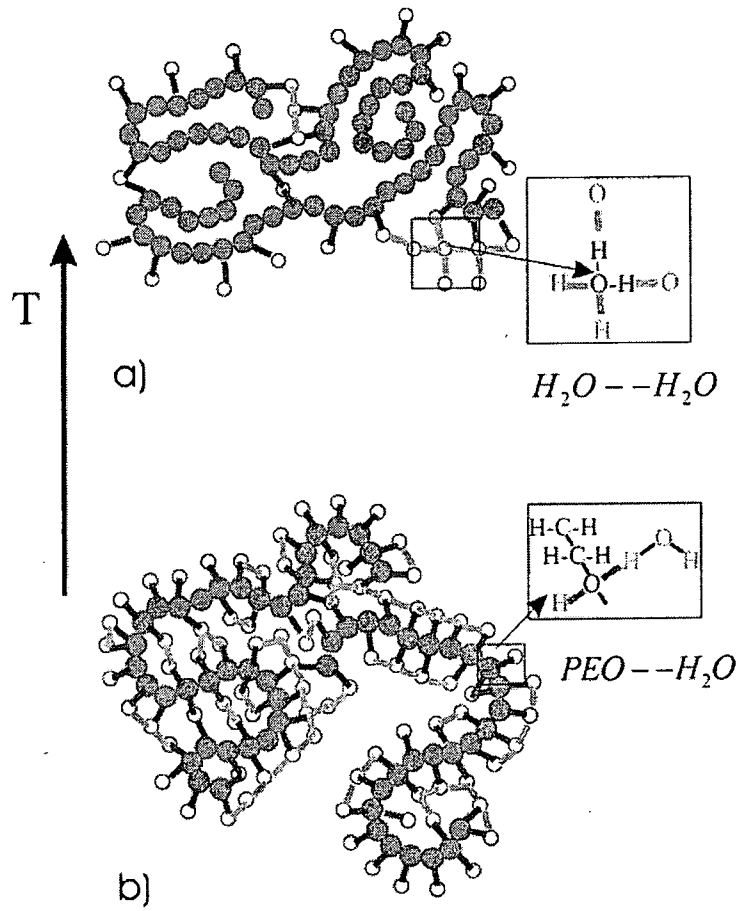


Figure 11 Schematic representation of hydrogen bonding between water and PEO, including water-water hydrogen bonds, and the corresponding chain conformation of poly(ethylene oxide) (PEO) in aqueous solutions at high (a) and low (b) temperatures.

ESEM studies were performed in order to investigate whether the drying process affected the aggregates' morphology (Figures 12-15).

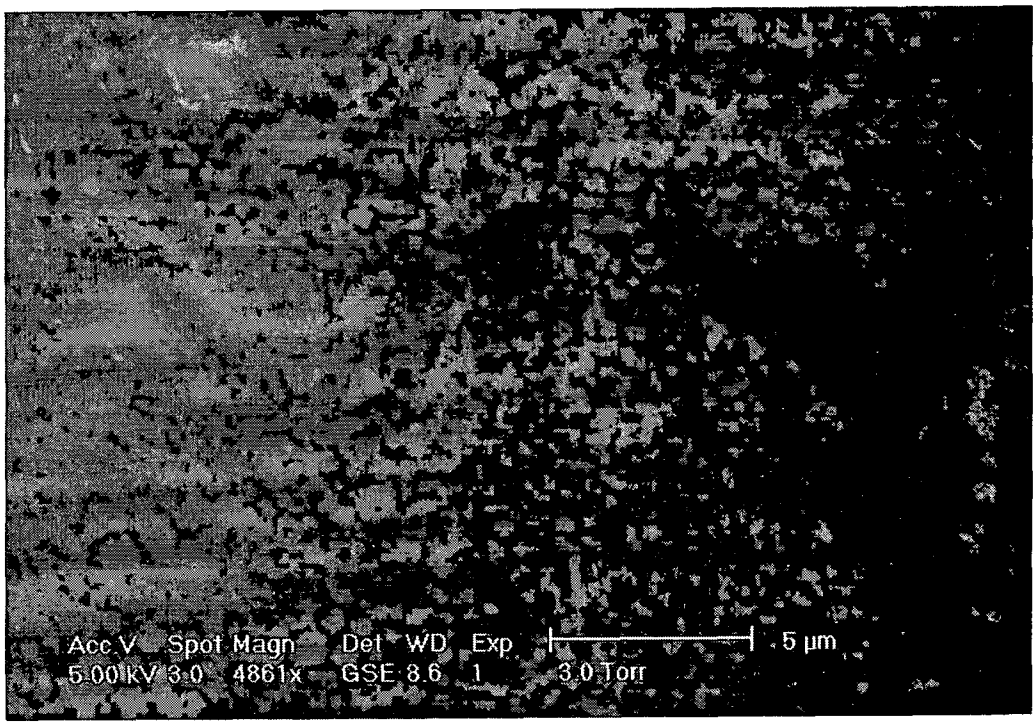


Figure 12 ESEM micrograph obtained using 5×10^{-2} mg ml^{-1} aqueous solution of pGalEMA-A, working at 3.0 Torr (wet mode).

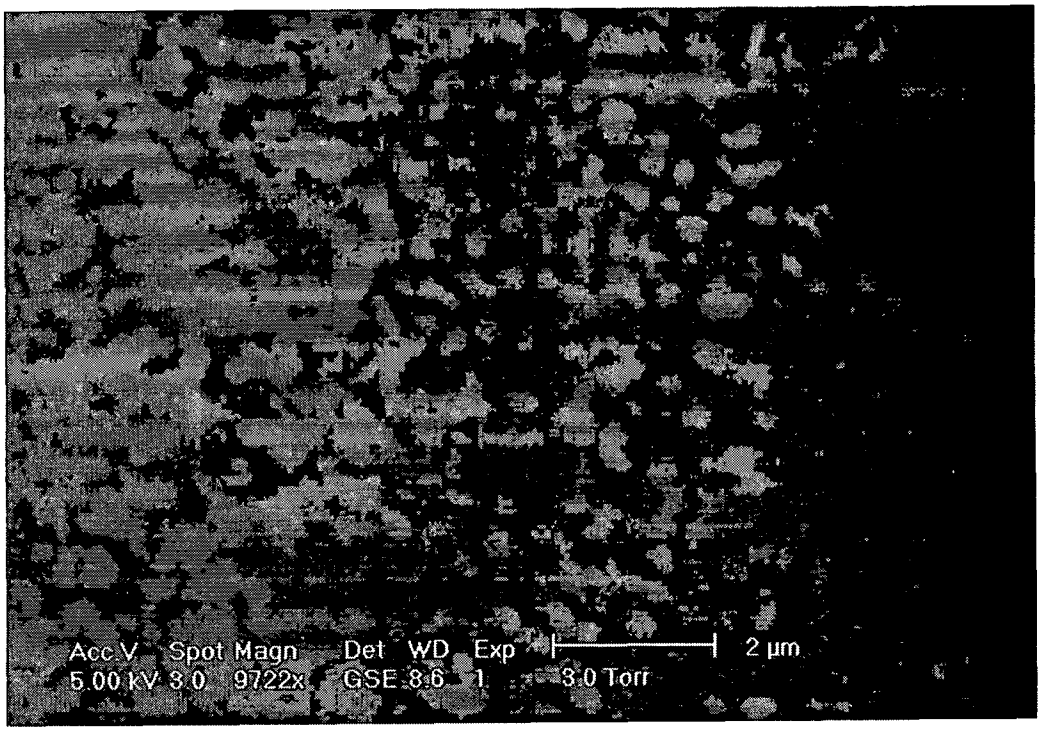


Figure 13 ESEM micrograph obtained using 5×10^{-2} mg ml^{-1} aqueous solution of pGalEMA-A, working at 3.0 Torr (wet mode).

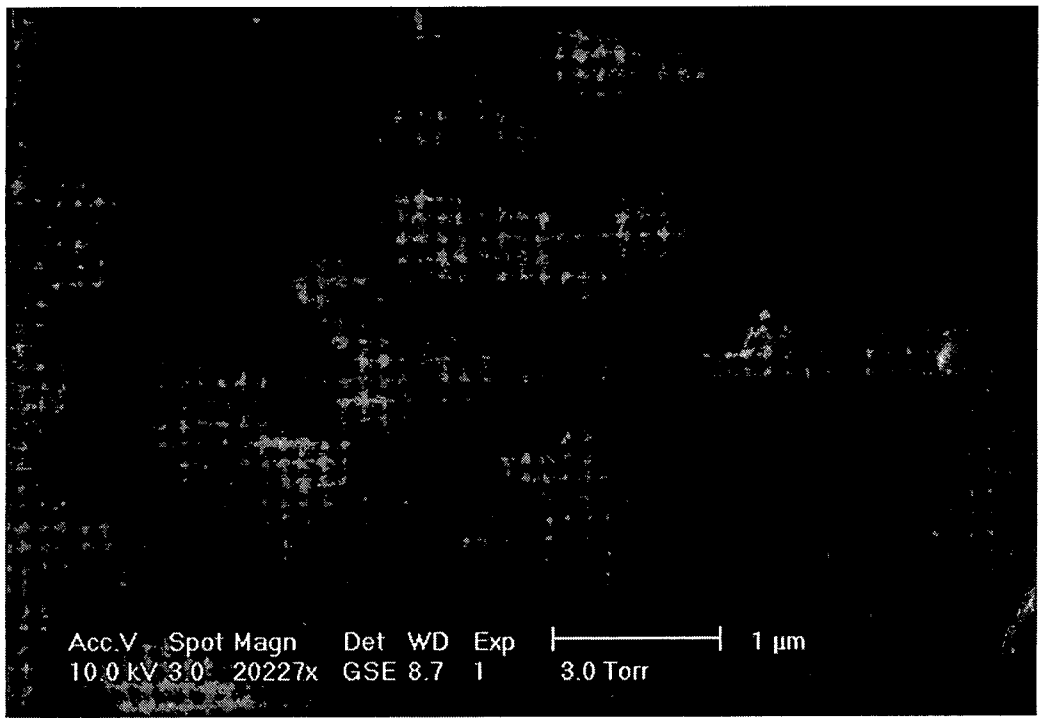


Figure 14 ESEM micrograph obtained using 5×10^{-2} mg ml^{-1} aqueous solution of pGalEMA-A, working at 3.0 Torr (wet mode).

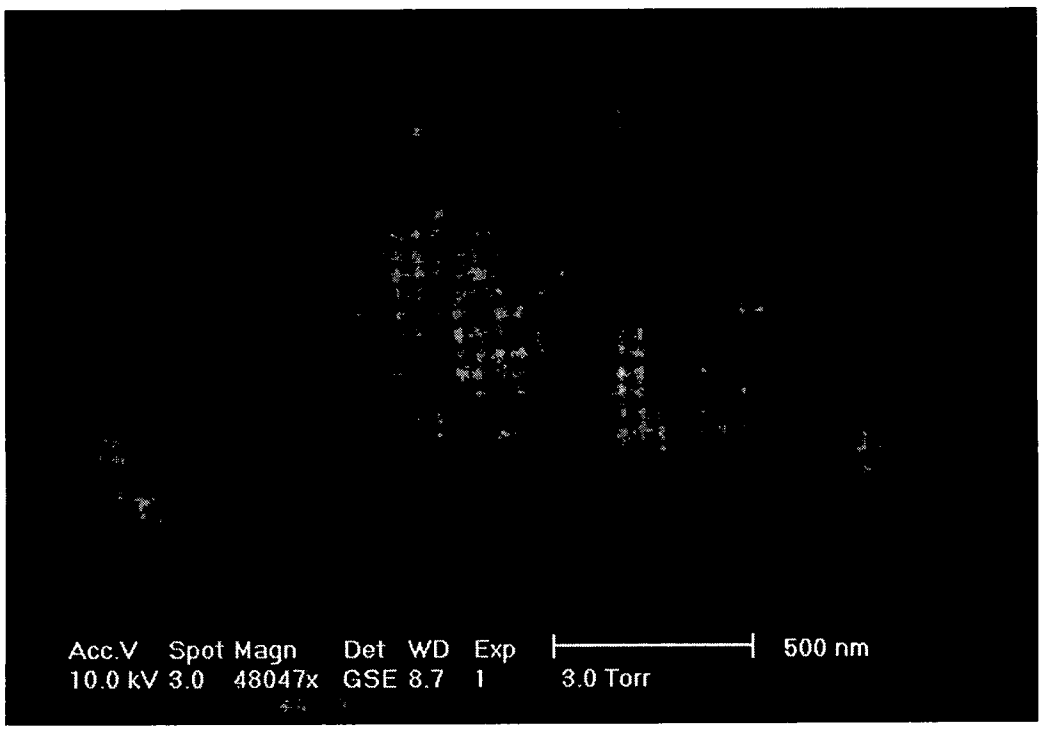


Figure 15 ESEM micrograph obtained using 5×10^{-2} mg ml^{-1} aqueous solution of pGalEMA-A, working at 3.0 Torr (wet mode).

A large amount of aggregates could be detected in solution, with diameter varying from approximately 50 to 500-600 nm, in good agreement with values determined by TEM. Unfortunately, it was not possible to obtain perfectly focused images of the particles in the wet mode and no conclusions could be drawn on the variation of the aggregates' morphology upon drying. Nevertheless, as reported in Figure 15, the results seemed supporting the hypothesis of loose aggregates.

4.2.2 LECTIN-LIGAND INTERACTION

4.2.2.1 ULTRAVIOLET-DIFFERENCE SPECTROSCOPY

The interaction of peanut agglutinin (PNA) with D-galactose, D-lactose, GalEMA and pGalEMA-A was qualitatively investigated by ultraviolet-difference spectroscopy. The spectra obtained are reported in Figure 16.

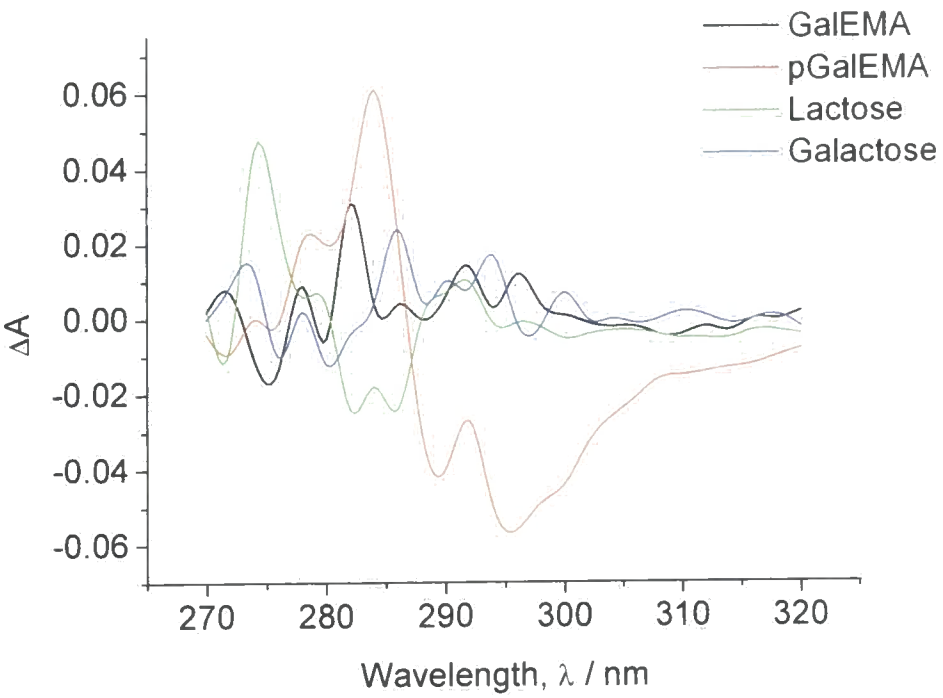


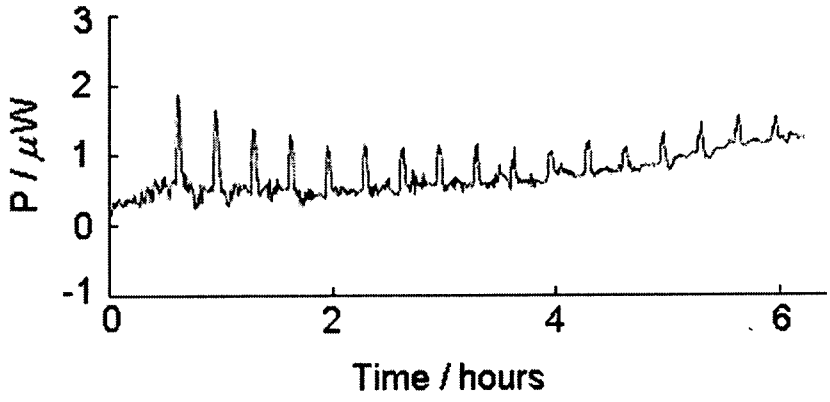
Figure 16 UV-difference spectra of peanut agglutinin with D-galactose, D-lactose, GalEMA and pGalEMA-A. [PNA] = 8×10^{-5} M, [ligands] = 2×10^{-4} M. The solutions were in 10 mM PBS at pH 7.4.

The presence of UV-difference spectra indicated that all ligands bound to the lectin. Interestingly, the intensity of the UV-difference spectrum increased in the order: D-galactose, GalEMA, D-lactose, pGalEMA-A. Moreover, the addition of GalEMA and pGalEMA to the lectin solution generated a UV-difference spectrum showing a maximum at approximately the same wavelength. GalEMA and pGalEMA contain the same β-D-galactosyl residue, whose position in the lectin binding site in PNA-GalEMA and PNA-pGalEMA complexes may be assumed to be similar. Although there is no demonstrated linear correlation between the intensity of the UV-difference spectrum and the ligand affinity, we were pleased to see broad agreement between the spectroscopic and calorimetric measurements (*vide infra*).

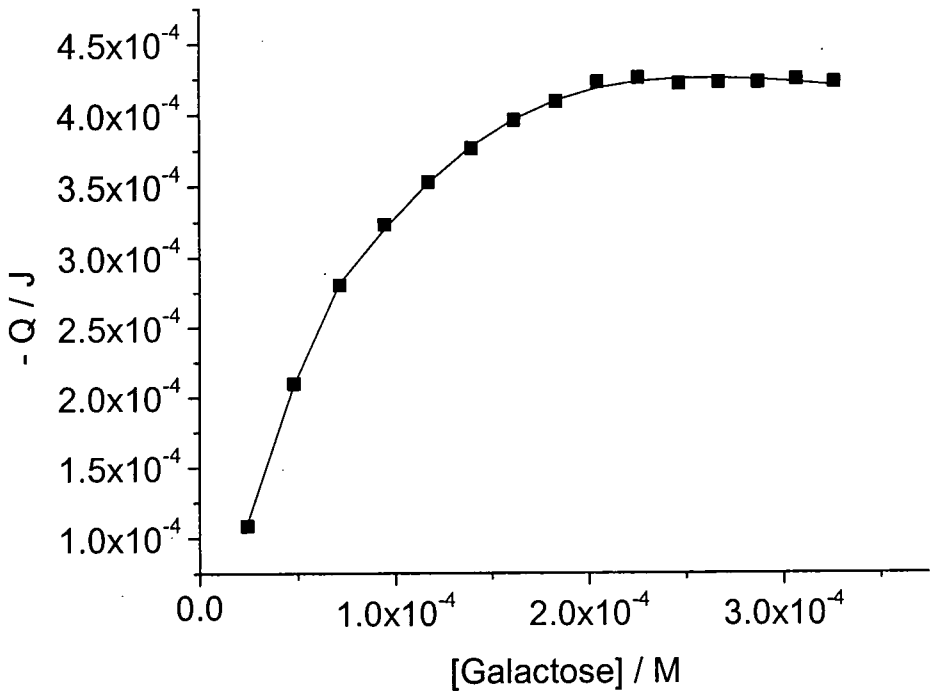
4.2.2.2 ISOTHERMAL TITRATION MICROCALORIMETRY

Once the binding was assessed by UV-difference spectroscopy, we proceeded to a quantitative determination of the thermodynamics by isothermal titration microcalorimetry. Calorimetric determinations of the binding affinity of D-galactose, GalEMA and pGalEMA-A,B for peanut agglutinin were performed. Reactant solutions were prepared in citrate buffer. Despite the limited solubility of PNA in this medium, citrate buffer was used in order to remove any possible contribution of buffer protonation effects to the enthalpic term. Citrate buffer, in fact, exhibits ΔH_{ion} = 0 kJ mol^{-1}[73]. As already mentioned, thermodynamic parameters are state functions and protein-ligand complexation, changes in protonation state and coupled equilibria are all reported as a single enthalpy value. Thermodynamic parameters were determined by non-linear least-squares analysis of the binding isotherms. Typical outputs of the raw microcalorimetry data, the binding isotherms and the best curve

fittings determined for PNA-D-galactose and PNA-pGalEMA-A interactions are shown in Figures 17 and 18, respectively.

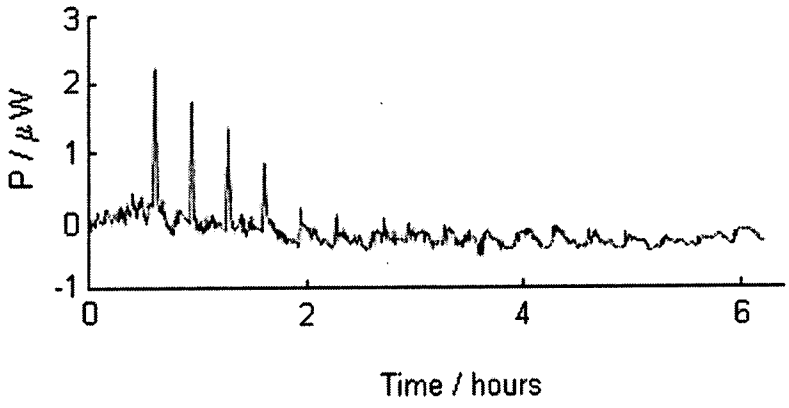


(a)

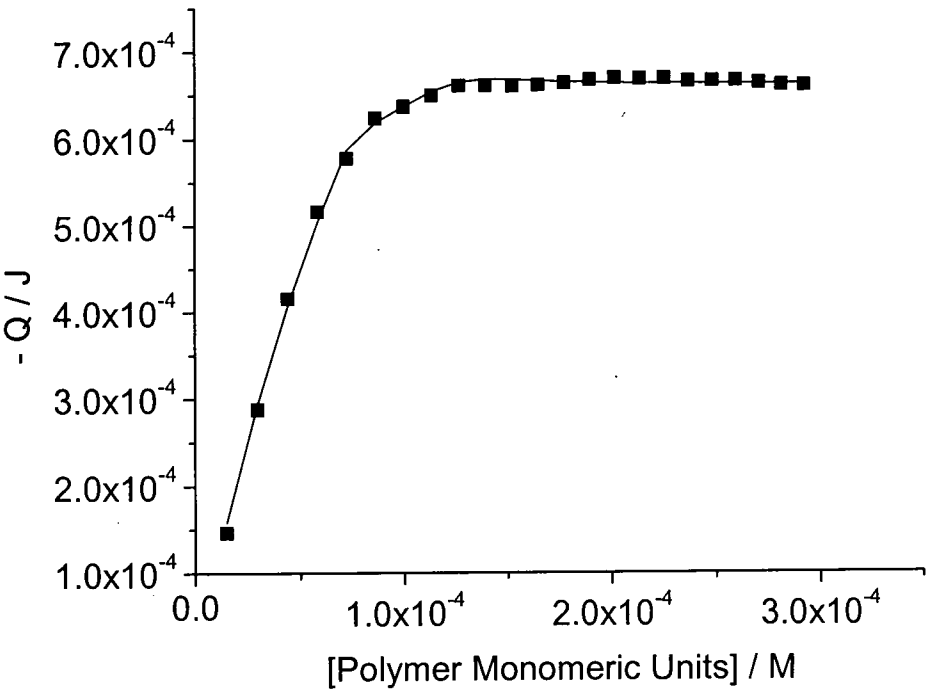


(b)

Figure 17 a), Experimental calorimetric data for the isothermal titration at 298 K of D-galactose into PNA. [PNA] = 3 mg ml^{-1}, [Ligand] = 0.55 mg ml^{-1}. The solutions were in 20 mM citrate buffer containing 150 mM NaCl, pH 7.4; b), corresponding binding isotherm and best fit curve to the data.



(a)



(b)

Figure 18 a), Experimental calorimetric data for the isothermal titration at 298 K of pGalEMA-A into PNA. [PNA] = 3 mg ml^{-1}, [Ligand] = 0.55 mg ml^{-1}. The solutions were in 20 mM citrate buffer containing 150 mM NaCl, pH 7.4; b), corresponding binding isotherm and best fit curve to the data.

The model used in this study for the non-linear least-squares analyses was firstly designed for the determination of the thermodynamics of the interaction of DNA with cationic polymers[74]. Its employment here was due to its unique property

of considering possible precipitation processes concomitant with or following the protein-ligand interaction. The nature of the model is critical for calorimetric analyses. In all cases the experimental data could not be fitted by curves generated by non-linear least-squared deconvolutions carried out using a generally utilised model for a 1:1 non-cooperative interaction where possible precipitation was not taken into account. The employment of the chosen model for the investigation of our protein-carbohydrate systems was supported not only by the good agreement between the experimental data and the curve fittings, but also by dynamic light scattering measurements. Moreover, we were pleased to see good agreement of the thermodynamic binding parameters determined for D-galactose with those reported in the literature for the same ligand[28].

For all ligands, DLS measurements showed an increase of the intensity of the scattered light throughout the titrations (Figures 19 and 20), indicating a precipitation process accompanying the protein-sugar interaction. For D-galactose and the monomeric ligand, GalEMA, no significant increase of scattering intensity was observed until sugar:lectin monomer ratios of *ca.* 1 (for both ligands when [PNA] = 1.5 mg ml^{-1}, in citrate buffer) and *ca.* 1.5-2 (for D-gal, [PNA] = 3 mg ml^{-1}, in saline citrate buffer) were reached. These values agreed with the interaction stoichiometry determined by ITM (*vide infra*). The results indicated that there was no significant precipitation at low ligand:lectin ratios; as the ratio increased and the interaction reached completion, precipitation occurred, probably due to a decrease of the already limited lectin solubility upon ligand binding. Indeed, the solutions, clear at the beginning of the experiments, became cloudy during the titrations.

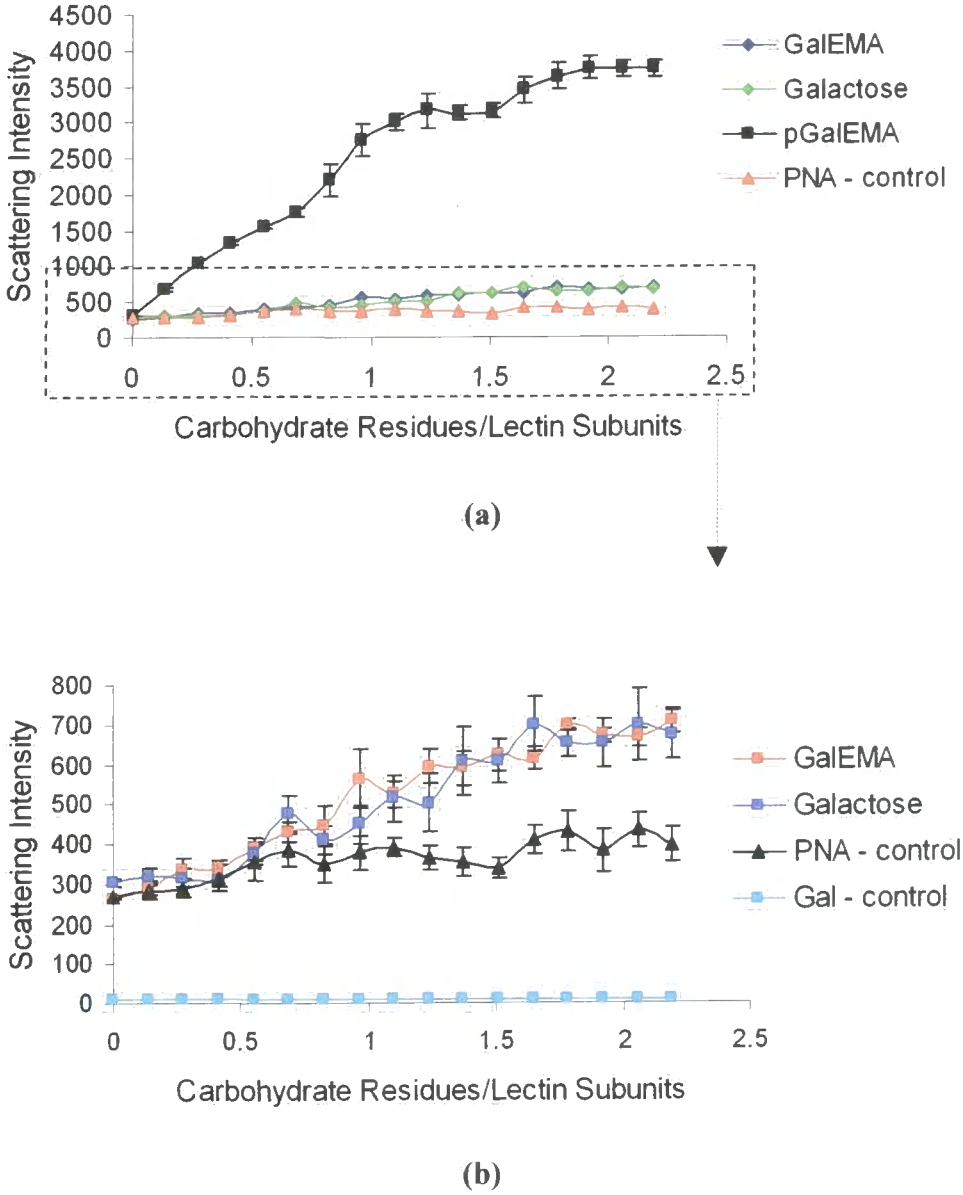
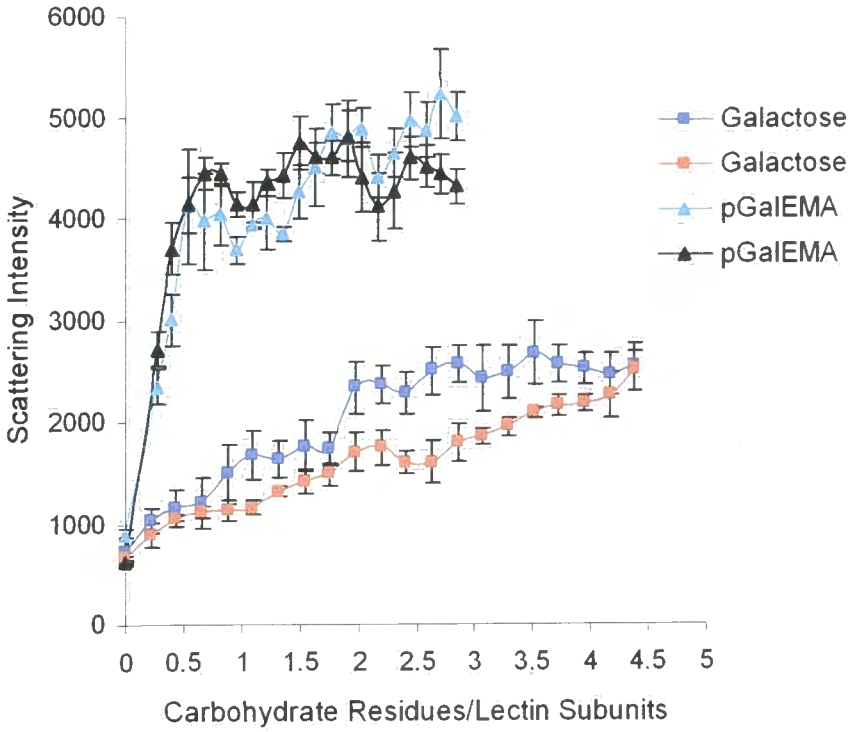
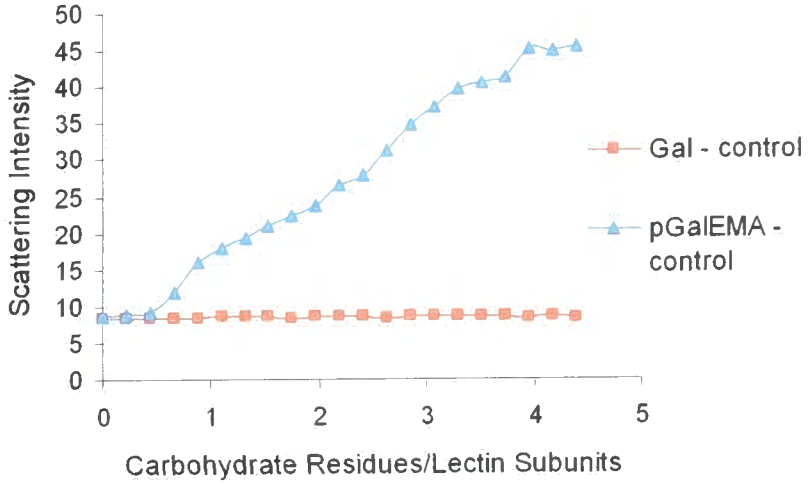


Figure 19 Scattering intensity measurement for PNA-ligand complexes' formation. [PNA] = 1.5 mg ml^{-1}. Each point represents the addition of aliquots of ligand solution, reproducing the titration experiments carried out by ITM. D-galactose was injected into buffer without lectin and buffer was injected into PNA as control. The experiments were carried out in 20 mM citrate buffer, pH 7.4 at 25 °C.



(a)



(b)

Figure 20 Scattering intensity measurement for PNA-ligand complexes' formation. a), Each point represents the addition of aliquots of ligand solution to PNA solution, reproducing the titration experiments carried out by ITM, [PNA] = 3 mg ml^{-1}; b), each point represents the addition of aliquots of ligand solution to buffer without lectin; the error bars are within the symbols. The experiments were carried out in 20 mM citrate buffer, containing 150 mM NaCl, pH 7.4 at 25 °C.

When pGalEMA-A was used as titrant, a significant increase of the intensity of scattered light was detected already after the first addition. The magnitude of the

increase was much larger than that recorded for the monovalent ligands. These results, combined with what has just been reported for the monomer, GalEMA, and with the thermodynamic binding parameters determined by ITM (*vide infra*), would suggest the formation of large, cross-linked PNA-polymer complexes, which phase-separate during the experiments. Formation of a precipitate was also observed by the naked eye. An intramolecular mechanism of protein-ligand interaction would result in the formation of a complex where intersite regions of the receptor surface are covered by polymer segments rich in hydrophilic groups, which would not be expected to lower the lectin solubility. Moreover, pGalEMA is highly soluble in aqueous media. Therefore, as a first approximation, the effect of pGalEMA chelation upon PNA solubility could be expected to be analogous to that determined for GalEMA if an intramolecular mechanism was operating, *i.e.* a lower increment of scattered light would be recorded. Again DLS measurements showed good agreement with calorimetric results: a plateau was reached at the same addition/injection for analogous experiments.

Calorimetrically determined thermodynamic binding parameters are reported in Table 4[§]. As expected monovalent ligands gave rise to enthalpically driven protein-ligand interactions, exhibiting enthalpy-entropy compensation[40,70]. The favourable enthalpic contribution arises from the formation of direct and water-mediated hydrogen bonds and van der Waals interactions, while the negative entropic term may be mainly attributed to the loss of degrees of freedom of the ligand upon binding[21]. The values of K and n for D-galactose were approximately 1×10^{-3} M^{-1} and 1, respectively, and showed good agreement with those previously reported[28].

[§] Data presented here are the mean of a minimum of three replicate titrations for each experiment. Since molar concentrations of the reactants rather than their activities were used, the subscript "*obs*" was affixed to the calculated binding parameters.

Table 4 Thermodynamic binding parameters of D-galactose, GalEMA and pGalEMA-A,B to peanut agglutinin.

EXPERIMENT – No.	n	$10^3 \times K_{obs}$ M^{-1}	ΔH^0_1 kJ mol^{-1}	ΔG^0_{obs} kJ mol^{-1}	$T\Delta S^0_{obs}$ kJ mol^{-1}	n_s	ΔH^0_2 kJ mol^{-1}
D-Galactose[a]	1.00 ± 0.42	0.95 ± 0.06	-22.97 ± 0.87	-16.98 ± 0.16	-5.99 ± 1.03	0.96 ± 0.35	-0.74 ± 0.31
GalEMA[a]	1.13 ± 0.04	5.65 ± 0.10	-39.41 ± 0.36	-21.41 ± 0.05	-18.00 ± 0.41	1.21 ± 0.16	-7.58 ± 0.10
pGalEMA-A[a]	1.56 ± 0.39	46.36 ± 8.81	-5.79 ± 0.19	-26.62 ± 0.48	20.83 ± 0.67	1.95 ± 0.1	-1.39 ± 0.01
pGalEMA B[a]	11.04 ± 0.50	9.40 ± 1.05	-19.84 ± 0.52	-22.67 ± 0.27	2.83 ± 0.79	9.19 ± 2.85	-0.93 ± 0.02
D-Galactose[b]	2.08 ± 0.22	0.78 ± 0.03	-23.94 ± 0.75	-16.51 ± 0.09	-7.43 ± 0.84	1.04 ± 0.04	-1.30 ± 0.03
pGalEMA-A[b]	1.62 ± 0.11	6.42 ± 0.37	-10.80 ± 0.19	-21.72 ± 0.14	10.92 ± 0.33	1.44 ± 0.17	-2.36 ± 0.007
pGalEMA-A[b]	1.35 ± 0.34	5.17 ± 0.79	-8.30 ± 0.29	-21.19 ± 0.38	12.89 ± 0.67	1.94 ± 0.58	-1.46 ± 0.002
pGalEMA-A[b*]	1.00 ± 0.36	4.17 ± 0.57	-6.42 ± 0.18	-20.65 ± 0.34	14.23 ± 0.52	1.85 ± 0.62	-1.20 ± 0.01
pGalEMA-A[c]	4.03 ± 0.08	1.37 ± 0.03	-15.76 ± 0.18	-17.90 ± 0.05	2.14 ± 0.23	4.99 ± 0.66	-1.53 ± 0.01
pGalEMA-A[c*]	4.07 ± 0.31	1.84 ± 0.14	-42.64 ± 2.57	-18.63 ± 0.19	-24.01 ± 2.76	2.99 ± 0.73	-4.56 ± 0.07

[a] [PNA] = 1.5 mg ml^{-1}, 20 mM citrate buffer, pH 7.4; [b] [PNA] = 3 mg ml^{-1}, 20 mM citrate buffer, containing 150 mM NaCl, pH 7.4; [b*] [PNA] = 3mg ml^{-1}, 20 mM citrate buffer, containing 150 mM NaCl, pH 7.4, heats of dilution have been subtracted; [c] [PNA] = 5.3 mg ml^{-1}, 50 mM Tris-HCl buffer, containing 500 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM $CaCl_2$, pH 6.9; [c*] [PNA] = 5.3 mg ml^{-1}, 50 mM Tris-HCl buffer, containing 500 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM $CaCl_2$, pH 6.9, heats of dilution have been subtracted.

GalEMA exhibited a 5-fold enhanced binding affinity compared to D-galactose, due to an increase of $-\Delta H^{0**}$ uncompensated by $-\Delta S^0$. This result is in agreement with the previously reported specificity of PNA for sugars in the pyranose form[48]. Significantly different thermodynamics were determined for polymeric ligands. On a valency-corrected basis, pGalEMA-A exhibited a 50-fold affinity gain compared to D-galactose and a 10-fold enhancement with respect to the corresponding monovalent ligand, GalEMA. A significant diminution of enthalpy was determined: $-\Delta H^0$ decreased from *ca.* -23 kJ mol^{-1} for D-galactose to *ca.* -6 kJ mol^{-1} for pGalEMA-A. On the contrary, the entropic contribution became favourable, varying from negative for the monovalent ligands to positive for the polymer. In particular, for pGalEMA-A approximately 75% of ΔG^0 arose from the favourable entropy term.

It is interesting to compare the results obtained for the two polymeric ligands pGalEMA-A and pGalEMA-B[††]. An affinity enhancement was still detected for pGalEMA-B compared to D-galactose and GalEMA: approximately 10-fold and 2-fold, respectively. With respect to the monovalent compounds, the polymer interaction was still characterised by a decrease of $-\Delta H^0$ and a favourable ΔS^0. However, its affinity for the protein was much lower than that of pGalEMA-A. The binding constant diminished by approximately 5-fold, due to the reduction of the favourable entropic term, $T\Delta S^0$, from *ca.* 20 kJ mol^{-1} for pGalEMA-A to *ca.* 3 kJ mol^{-1} for pGalEMA-B. $-\Delta H^0$, instead, raised to *ca.* -20 kJ mol^{-1}, giving a more favourable contribution than that reported for pGalEMA-A. The stoichiometries observed for the

---

** ΔH^0_1, ΔH^0_2, n and n_s are all model parameters. The results are discussed in terms of ΔH^0_1 and n. ΔH^0_2 was less significant and mostly invariant regardless the ligand type. Furthermore, ΔH^0_1 and ΔH^0_2 are mutually related by model assumption; therefore, changes in ΔH^0_2 are reflected in ΔH^0_1 values. n_s was approximately equal to n for all experiments.

†† pGalEMA-A: M_n = 461,000 g mol^{-1} with M_w/M_n = 2.21; pGalEMA-B: M_n = 2,984 g mol^{-1} with M_w/M_n = 2.15.

protein-carbohydrate interactions were also intriguingly different, being 1.5 and 11 for pGalEMA-A and pGalEMA-B, respectively. As mentioned above, the polymer deprotective step involved in the synthesis of pGalEMA-B is non-quantitative, resulting in a polymeric material of ill-defined composition, where partially or fully protected sugar residues, unable to bind (efficiently) to the lectin, are still present. Thus, the lower affinity of pGalEMA-B relative to pGalEMA-A and the higher number of carbohydrate residues per lectin monomer required for the interaction can easily be understood, considering that not all the sugar residues can, in principle, be recognised by the protein. This confirmed our expectations that fully deprotected polymers and partially acetylated ones would possess different functional properties. Importantly, it unambiguously shows how strongly the polymer binding properties are affected by its composition, leading again to the indispensability of highly pure and strictly well-defined materials.

According to Toone and co-workers[11,16], a decrease of the binding enthalpy (in absolute value) can be considered as the thermodynamic signature of an endothermic intermolecular aggregative process superimposed on an exothermic ligand binding event, which is assumed to proceed with thermodynamic parameters equivalent to those of the corresponding monovalent saccharide[12]. Therefore, the thermodynamic parameters evaluated for polymeric ligands in this study, and especially those for pGalEMA-A, can be related to an affinity enhancement mainly achieved through the formation of cross-linked ligand-protein complexes[12,75,76], which phase separate during the titration. A significant amount of precipitate was always observed in the cell at the end of the calorimetric measurements. In analogy with the hydrophobic effect responsible for the micellisation of surfactants in aqueous media, the significant favourable entropic contribution might then be explained by the

release of structurally ordered water molecules from the hydration shells of protein tetramers brought close to one another by the aggregation (each peanut agglutinin tetramer is surrounded by 518 molecules of water[50]). This hypothesis was also supported by dynamic light scattering measurements.

At least two phenomena have to be taken into account in the search for a justification of the determined protein-pGalEMA-A binding stoichiometry of approximately 1.5. First of all, the proximity of carbohydrate residues to each other within the polymer might prevent the interaction of all the sugars, due to steric hindrance. Secondly, the observed precipitation of the cross-linked aggregates would decrease the number of binding sites available in solution. Lectin-carbohydrate cross-linked complexes have been found to generally occur during protein-multivalent ligand interactions[12,18,75,76]. Their formation was found to depend on several factors, including the valencies of the two interacting molecules, the structural arrangement of the ligand epitopes, the binding sites and the subunits of the lectin. When both valencies are greater than two, 2- and 3-dimensinal cross-linking was reported, giving rise to highly ordered complexes.

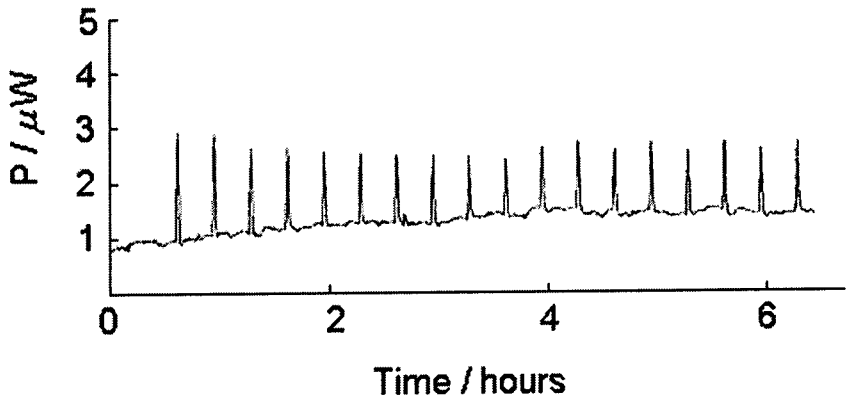
Rao *et al.*[22,23] recently measured the thermodynamics of the interaction of a trivalent derivative of DADA (D-Ala-D-Ala) and of Lac-R'$_d$-Lac with a trivalent and a divalent derivative of vancomycin, respectively. The observed affinity enhancements compared to the monovalent analogues were in both cases certainly reached by means of an intramolecular mechanism. Both $-\Delta H$ and $-\Delta S$ were found to scale proportionally to the number of epitopes, behaviour not observed in this study.

Despite glycopolymers possessing a carbohydrate residue per monomeric unit have been shown to bind lectins through a chelation mechanism[77], the thermodynamic data, the stoichiometry of the interaction, DLS measurements and the

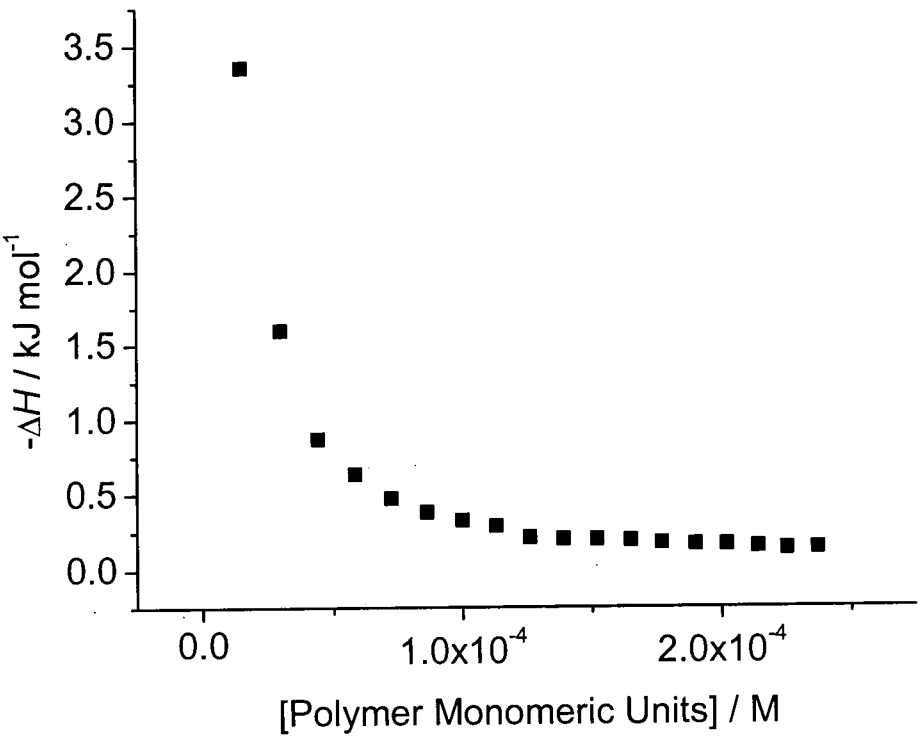
lack in the polymeric structure of portions resembling the protein surface, which could lead to favourable interactions in a chelation process, indicate that the polymers analysed in this study are likely to act mainly by aggregating distinct PNA tetramers[16].

Addition of salt slightly affected the interaction of D-galactose with the lectin. Only a small decrease of the binding constant was evaluated with a slight increase of the sugar:lectin subunit ratio. Binding isotherms became less steep as a result of the diminished affinity strength. On the contrary, the presence of NaCl critically affected the interactions of GalEMA and pGalEMA. Calorimetric studies of PNA-GalEMA binding under these conditions were no longer possible, probably due to the too low [PNA monomer]·K product[78]. The binding constant of pGalEMA-A decreased to *ca.* 6×10^{-3} M^{-1}, approximately 7-fold lower than that determined in the absence of salt. The enthalpy change was not strongly affected, becoming slightly more favourable, while $T\Delta S^0$ almost halved. A lower incidence of the endothermic, entropically driven, aggregation process would result in a more favourable enthalpic term and a decrease of the positive entropic contribution. A further increase in the amount of salt still further reduced the binding constant of PNA-polymer interaction to *ca.* 1×10^{-3} M^{-1}, as shown by experiments carried out in 50 mM Tris-HCl buffer containing 500 mM $MgCl_2$, 1 mM $CaCl_2$ and 1 mM $MnCl_2$. Although Tris-HCl exhibits ΔH_{ion} = 46.02 kJ mol^{-1} and buffer protonation effects may affect ΔH and ΔS values, favourable enthalpic and entropic contributions were still observed. An increase of n (which is not buffer dependent) up to 4 was determined. A weaker interaction would require an excess of polymer to reach protein binding sites' saturation.

It is interesting to note that blank experiments carried out injecting the polymer into citrate and Tris-HCl buffers gave significantly different results (Figures 21 and 22).

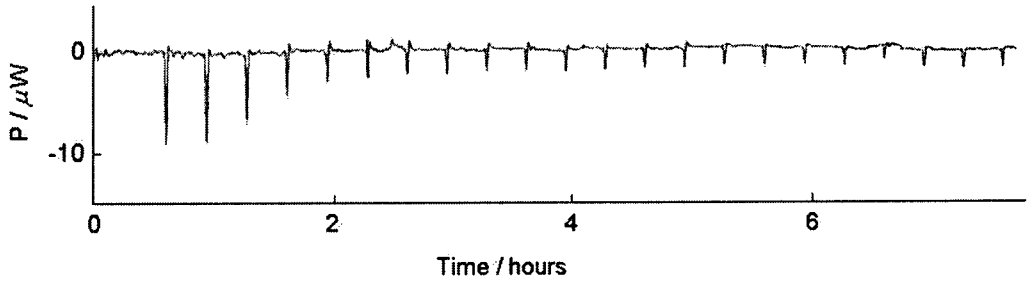


(a)

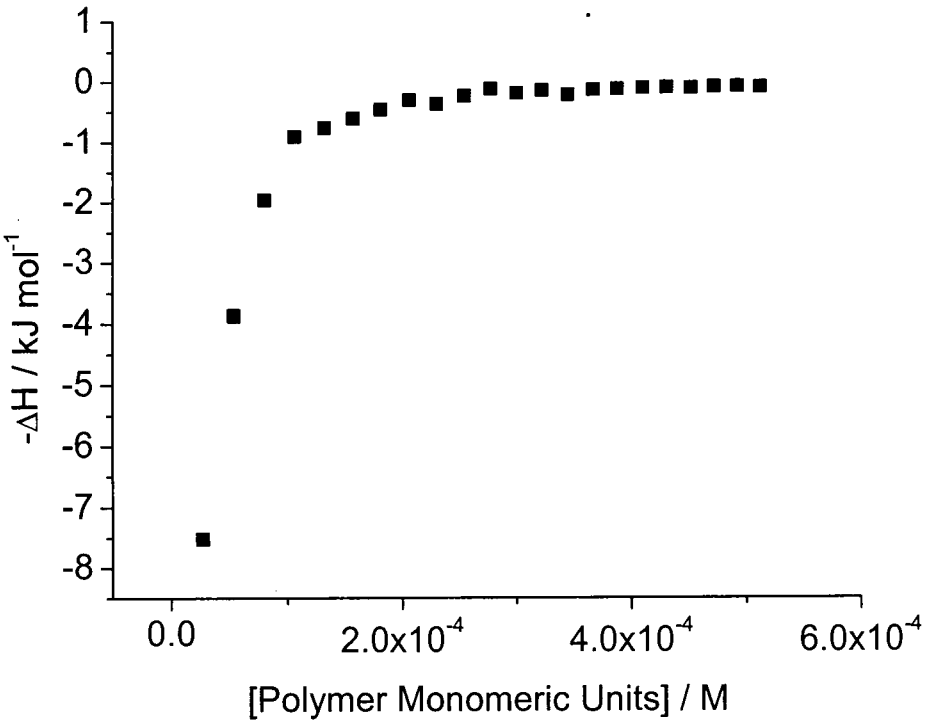


(b)

Figure 21 a), Experimental calorimetric data for the injection of pGalEMA-A into buffer only. [polymer] = 0.55 mg ml^{-1}, 20 mM citrate buffer, containing 150 mM NaCl, pH 7.4, 25 °C; b), corresponding calculated isotherms -ΔH *versus* polymer effective concentration.



(a)



(b)

Figure 22 a), Experimental calorimetric data for the injection of pGalEMA-A into buffer only. [polymer] = 1 mg ml^{-1}, 50 mM Tris-HCl buffer, containing 500 mM $MgCl_2$, 1 mM $CaCl_2$ and 1 mM $MnCl_2$, pH 6.9, 25 °C; b), corresponding calculated isotherms -ΔH *versus* polymer effective concentration.

In both cases non-negligible heats of dilution were observed. It has been already discussed that DLS, TEM and ESEM measurements showed that deprotected glycopolymers tended to self-associate in water. DLS studies revealed that pGalEMA formed nanoaggregates also in citrate and Tris-HCl buffer solutions. Whilst

aggregates were present in the starting polymer solutions used for the ITM titrations, no trace of aggregation could be detected at concentrations as low as those in the microcalorimeter cell after the first injection. Thus, the heats of dilution observed by ITM could be related to the dissociation of the aggregates under a certain critical concentration. Despite the fact that a full calorimetric study of the polymer behaviour in aqueous media was not carried out, it is worth noting that good agreement between the critical aggregation concentration values in the two buffer solutions determined by DLS and estimated by ITM was observed. A change of medium from water to saline citrate buffer seemed not to affect significantly the polymer behaviour in solution. In both cases CAC was close to 1×10^{-2} mg ml^{-1}. In Tris-HCl, instead, pGalEMA-A aggregation appeared to occur at a higher concentration approximately equal to 4×10^{-2} mg ml^{-1}, as determined by both techniques. In all media, the average diameter of the aggregates was close to 40 nm when the polymer concentration was *ca.* 1 mg ml^{-1}. DLS measurements revealed that the aggregate dimension increased with increasing concentration. For experiments carried out in saline citrate buffer, this further association could be detected by microcalorimetry as an exothermic process (the intensity of the peaks reported in Figure 21a decreases until the CAC is reached, well above this concentration the intensity of the peaks increases due to the further association of the aggregates present in solution).

It is worth noting that when the calorimetric experiments were performed in citrate buffer the dissociation of the polymer aggregates was exothermic, while the same was endothermic in Tris-HCl. Buffer protonation was therefore certainly involved in this deaggregation process. The exchange of protons between the buffer and the polymer might result in electrostatic interactions somehow involved in the binding of pGalEMA to PNA. Moreover, ions resulting from the salt dissolution

would neutralise charges on the protein surface involved in intramolecular and intermolecular electrostatic interactions. Thus, a change in the quaternary structure of the protein could affect the chelation of different binding sites by the polymer. On the other hand, protein-protein attractive interactions, perhaps involved in the aggregation of tetramers cross-linked by the ligand, could be neutralised hampering the aggregation itself.

It is worthwhile to note that for experiments performed in citrate buffer thermodynamic data determined by subtracting heats of dilution (evaluated by means of blank titrations) from each binding heat or by subtracting the average dilution heat determined from the end part of the release heat titration profile[79] were substantially identical. Only a slight decrease of ΔG^0 was observed, due to a slight decrease of the enthalpic term (Table 3). Therefore, the concomitancy of the polymer aggregates' dissociation process with the binding event did not prevent the determination of thermodynamic binding parameters.

4.3 EXPERIMENTAL

4.3.1 DYNAMIC LIGHT SCATTERING, DLS

Dynamic light scattering (DLS) measurements were performed using a Brookhaven RI 200SM goniometer and a BI-9000AT digital correlator, with a frequency doubled solid state Nd-YAG laser (λ_0 = 532 nm) (for the determination of the critical aggregation concentration of pGlcEMA-A,B and pGalEMA-A,B in water and buffer solutions and for the determination of the aggregates' size distribution) and with a Malvern PCS4700 photon correlation spectroscopy assembly (Malvern Instruments, Marlvern, UK) (for the PNA-ligand complexes' formation). Titrations

analogous to those carried out by ITM were performed adding aliquots of ligand solution to the lectin solution placed into the cell. The intensity of the scattered light was measured after each addition of the titrant and reported *versus* carbohydrate residues:lectin subunits ratio. Control experiments adding ligand to buffer and buffer to PNA were also performed.

4.3.2 TRANSMISSION ELECTRON MICROSCOPY, TEM

Transmission electron microscopy (TEM) measurements were performed using a Philips CM100 (compustage) transmission electron microscope operating at 100 kV. One drop of a 10 $mg \cdot ml^{-1}$ sample solution was put onto a coated copper grid and allowed to dry before analysis.

4.3.3 ENVIRONMENTAL SCANNING ELECTRON MICROSCOPY, ESEM

All ESEM experiments a were carried out by a FEI XL-30 field emission gun instrument, with a Peltier sample stage and gaseous secondary electron detector (GSED), to produce an image. All samples were contained within a circular brass stub with a diameter of 1 cm and a depth of 0.5 cm. The temperature of the stub, and hence sample, was controlled using the Peltier cooling stage with a water/propanol coolant maintained at 283 K. A few drops of the undiluted sample were placed on the stub, left for a few minutes to equilibrate, before initiating pump down of the chamber from ambient to a few Torr. Images from the CCD camera confirmed that the sample remained stable, i.e. there was no 'boiling' of water during this pump-down procedure as the method adopted (pumping down in steps) ensured that a high percentage of water vapor remained in the sample chamber/environment. In all cases, the sample

was cooled to ~ 275 K (2 °C) and surrounded by water vapor at a pressure of *ca.* 3 Torr. Working distances were typically 8-10 mm and images were collected using an accelerating voltage of 5 keV. Under such conditions signal to noise ratio was maximized, sample dehydration minimized, and detailed high resolution images were obtained. Rod dimensions and their separation were subsequently determined directly from the images using computer software built into the ESEM.

4.3.4 ULTRAVIOLET-DIFFERENCE SPECTROSCOPY

Ultraviolet difference spectra were recorded on a Unicam UV2-100 spectrometer, using masked semimicro cells of 1-cm path-length. Aliquots of 500 μl of peanut agglutinin (PNA, Sigma, affinity-purified, salt-free lyophilised powder, activity < 0.1 μg ml^{-1}) solutions (*ca.* 2 mg ml^{-1}, *ca.* 80 μM when 27,000 g mol^{-1} was used as the molecular weight of PNA monomer, in PBS at pH 7.4) were added to both sample and reference cuvettes, and the baseline recorded. An aliquot (4 μl) of a 2.5×10^{-2} M solution of the ligand was added to the sample cell, while the reference received the same amount of buffer.

4.3.5 ISOTHERMAL TITRATION MICROCALORIMETRY

Isothermal titration microcalorimetry (ITM) was performed using a Thermal Activity Monitor (TAM 2277, Thermometric AB, Sweden) operated at 25 °C. Two sets of experiments were carried out preparing samples of PNA and ligands in 20 mM citrate buffer, containing or not 150 mM NaCl. The samples were adjusted to pH 7.4 (±0.05) using dilute NaOH. The addition of salt allowed the doubling of the lectin concentration in solution. In each experiment a 2.5 ml sample of a solution of PNA

(1.5-3 mg ml^{-1}) was placed in a sample cell and inserted into the instrument. Once thermal equilibrium was reached, the titration was performed by consecutive injections (25 injections of 20 μl each) of ligand solution (0.275-0.55 mg ml^{-1}). The titrant was added by means of a Hamilton microlab syringe mounted on a computer-operated syringe driver (Lund 6100 syringe pump). The experimental method set up via the Digitam^R 4 software allowed for data collection over a 15 minute period for the injection and a 5 minute baseline period before the next injection. This was found to be adequate for the interaction to proceed to completion at each injection and reach the baseline before the next addition. Data presented here are the mean of a minimum of three replicate titrations for each experiment. Heats of dilution/mixing were determined in blank titrations injecting aliquots of ligand solution into the appropriate buffer without the lectin. The signs of the released heat values were reversed for data analysis since the output from the instrument is from the perspective of the equipment and not the system under study. The equilibrium binding parameters K, ΔH^0 and n were determined by non-linear least-squares fitting using the routines available within Origin 5.0 (MicroCal Software). The integrated heats of binding were corrected for the enthalpy of dilution, ΔH_d, before calculation of the binding parameters. The heat of dilution can be evaluated either by performing a blank titration, *i.e.* injection of ligand solution into buffer without PNA, or from the end part of the released heat titration profile. In the latter case an average dilution heat is calculated and subtracted from the observed heat value at each data point[79]. In this study both methods were employed. When pGalEMA was used as the titrant, c values – [PNA monomer]K – were ≥ 0.9[78]. Concentrations of polymer and PNA are both expressed in terms of monomeric units. The free energy of complex formation (ΔG^0_{obs}) was calculated from $\Delta G^0_{obs} = -RT \ln K_{obs}$ and the entropy of binding determined as $T\Delta S^0_{obs} = \Delta H^0 - \Delta G^0_{obs}$.

4.4 CONCLUSIONS

Deprotected glycopolymers carrying glucose and galactose residues were found to self-associate in aqueous media, forming nanoaggregates with dimensions varying over a broad range of 15-600 nm. Considering the structure of the polymers and the results of fluorescence studies, it is very likely that the formed particles can be envisaged as loose aggregates, resulting from inter- and intramolecular direct and water-mediated hydrogen bonds.

Ultraviolet-difference spectroscopy showed that the monomer and polymer bearing galactose units (GalEMA and pGalEMA-A,B, respectively) were efficiently recognised and bounded by the galactose-specific lectin peanut agglutinin.

Quantitative evaluations of the thermodynamic binding parameters for the interaction of peanut agglutinin with D-galactose, GalEMA and pGalEMA-A,B were performed by isothermal titration microcalorimetry. Reported here is the first calorimetric study of the interaction of a lectin with polymeric ligands. A 50-fold affinity enhancement of the polymer compared to the monovalent saccharide D-galactose was observed. The thermodynamics and DLS measurements suggested that the polyvalent carbohydrates bound to the lectin mainly by cross-linking distinct protein molecules. The polymer activity was strongly influenced by the presence of salt, indicating that electrostatic interactions were involved in PNA-polymer complex formation, either by affecting the polymer nanoparticles dissociation, and/or neutralising charges on the protein surface somehow involved in the chelate/aggregative process. Inclusion in the polymer structure of segments resembling the protein surface, which could give rise to favourable interactions between the ligands and the receptor, as well as consideration of the target structure in the design of the ligands, would very likely improve the affinity gain.

4.5 BIBLIOGRAPHICAL REFERENCES

[1] N. Sharon, H. Lis, *Science* **1989**, *246*, 227-234.

[2] N. Sharon, H. Lis, *Lectins*, Chapman and Hall, London, **1989**.

[3] M. Takeichi, *Science* **1991**, *251*, 1451-1455.

[4] A. Varki, *Glycobiology* **1993**, *3*, 97-130.

[5] W. J. Lees, A. Spaltenstein, W. J. E. Kingery, G. M. Whitesides, *J. Med. Chem.* **1994**, *37*, 3419-3433.

[6] H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637-674.

[7] M. Mammen, S.-K. Choi, G. M. Whitesides, *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2754-2794.

[8] D. K. Mandal, N. Kishore, C. F. Brewer, *Biochemistry* **1994**, *33*, 1149-1156.

[9] Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lönngren, J. Arnarp, M. Haraldsson, H. Lönn, *J. Biol. Chem.* **1983**, *258*, 199-202.

[10] Y. C. Lee, in *Neoglycoconjugates: Preparation and Applications* (Eds.: Y. C. Lee, R. T. Lee), Academic Press, London, **1994**, pp. 3-21.

[11] J. J. Lundquist, E. J. Toone, *Chem. Rev.* **2002**, *102*, 555-578.

[12] S. M. Dimick, S. C. Powell, S. A. McMahon, D. N. Moothoo, J. H. Naismith, E. J. Toone, *J. Am. Chem. Soc.* **1999**, *121*, 10286-10296.

[13] G. Molema, D. K. F. Meijer, *Adv. Drug Deliv. Rev.* **1994**, *14*, 25-50.

[14] B. G. Davis, *Chem. Rev.* **2002**, *102*, 579-601.

[15] R. V. Weatherman, K. H. Mortell, M. Chervenak, L. L. Kiessling, E. J. Toone, *Biochemistry* **1996**, *35*, 3619-3624.

[16] J. B. Corbell, J. J. Lundquist, E. J. Toone, *Tetrahedron: Asymmetry* **2000**, *11*, 95-111.

[17] J. J. Lundquist, S. D. Debenham, E. J. Toone, *J. Org. Chem.* **2000**, *65*, 8245-8250.

[18] T. K. Dam, R. Roy, S. K. Das, S. Oscarson, C. F. Brewer, *J. Biol. Chem.* **2000**, *275*, 14223-14230.

[19] T. K. Dam, R. Roy, D. Pagé, C. F. Brewer, *Biochemistry* **2002**, *41*, 1351-1358.

[20] T. K. Dam, R. Roy, D. Pagé, C. F. Brewer, *Biochemistry* **2002**, *41*, 1359-1363.

[21] T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387-429.

[22] J. Rao, J. Lahiri, L. Isaacs, R. M. Weis, G. M. Whitesides, *Science* **1998**, *280*, 708-711.

[23] J. Rao, L. Yan, J. Lahiri, G. M. Whitesides, R. M. Weis, H. S. Warren, *Chem. Biol.* **1999**, *6*, 353-359.

[24] N. M. Young, A. Z. Johnston, D. C. Watson, *Eur. J. Biochem.* **1991**, *196*, 631-637.

[25] W. W. Fish, L. M. Hamlin, R. L. Miller, *Arch. Biochem. Biophys.* **1978**, *190*, 693-698.

[26] K. J. Neurohr, N. M. Young, H. H. Mantsch, *J. Biol. Chem.* **1980**, *255*, 9205-9209.

[27] K. J. Neurohr, D. R. Bundle, N. M. Young, H. H. Mantsch, *Eur. J. Biochem.* **1982**, *123*, 305-310.

[28] M. Caron, J. Ohanessian, J. Becquart, H. Gillier-Pandraud, *Biochim. Biophys. Acta* **1982**, *717*, 432-438.

[29] M. J. Blandamer, P. M. Cullis, J. B. F. N. Engberts, *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 2261-2267.

[30] M. Eftink, R. Biltonen, in *Biological Microcalorimetry* (Ed.: A. E. Beezer), Academic Press, London, **1980**, pp. 313-412.

[31] T. Lundback, H. Hansson, S. Knapp, R. Ladenstein, T. Hard, *J. Mol. Biol.* **1998**, *276*, 775-786.

[32] I. Jelesarov, H. R. Bosshard, *Biochemistry* **1994**, *33*, 13321-13327.

[33] G. C. Kresheck, L. B. Vitello, J. E. Erman, *Biochemistry* **1995**, *34*, 8398-8405.

[34] D. K. Srivastava, S. Wang, K. L. Peterson, *Biochemistry* **1997**, *36*, 6359-6366.

[35] D. A. Brummell, V. P. Sharma, N. N. Anand, D. Bilous, G. Dubuc, J. Michniewicz, C. R. MacKenzie, J. Sadowska, B. W. Sigurskjold, B. Sinnott, N. M. Young, D. R. Bundle, S. A. Narang, *Biochemistry* **1993**, *32*, 1180-1187.

[36] R. F. Kelley, M. P. O'Connell, *Biochemistry* **1993**, *32*, 6828-6835.

[37] D. R. Bundle, B. W. Sigurskjold, *Methods Enzymol.* **1994**, *247*, 288-305.

[38] C. Zentz, J.-P. Frénoy, R. Bourrillon, *FEBS Lett.* **1977**, *81*, 23-27.

[39] F. P. Schwarz, K. Puri, A. Surolia, *J. Biol. Chem.* **1991**, *266*, 24344-24350.

[40] B. A. Williams, M. C. Chervenak, E. J. Toone, *J. Biol. Chem.* **1992**, *267*, 22907-22911.

[41] F. P. Schwarz, K. D. Puri, R. G. Bhat, A. Surolia, *J. Biol. Chem.* **1993**, *268*, 7668-7677.

[42] D. K. Mandal, L. Bhattacharyya, S. H. Koenig, R. D. I. Brown, S. Oscarson, C. F. Brewer, *Biochemistry* **1994**, *33*, 1157-1162.

[43] M. C. Chervenak, E. J. Toone, *Biochemistry* **1995**, *34*, 5685-5695.

[44] D. Gupta, M. Cho, M. D. Cummings, C. F. Brewer, *Biochemistry* **1996**, *35*, 15236-15243.

[45] A. Surolia, N. Sharon, F. P. Schwarz, *J. Biol. Chem.* **1996**, *271*, 17697-17703.

[46] F. P. Schwarz, H. Ahmed, M. A. Bianchet, L. M. Amzel, G. R. Vasta, *Biochemistry* **1998**, *37*, 5867-5877.

[47] G. B. Reddy, V. R. Srinivas, N. Ahmad, A. Surolia, *J. Biol. Chem.* **1999**, *274*, 4500-4503.

[48] R. Lotan, E. Skutelsky, D. Danon, N. Sharon, *J. Biol. Chem.* **1975**, *250*, 8518-8523.

[49] N. M. Young, R. P. Oomen, *J. Mol. Biol.* **1992**, *228*, 924-934.

[50] R. Banerjee, K. Das, R. Ravishankar, K. Suguna, A. Surolia, M. Vijayan, *J. Mol. Biol.* **1996**, *259*, 281-296.

[51] R. Banerjee, S. C. Mande, V. Ganesh, K. Das, V. Dhanaraj, S. K. Mahanta, K. Suguna, A. Surolia, M. Vijayan, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 227-231.

[52] V. R. Srinivas, G. B. Reddy, N. Ahmad, C. P. Swaminathan, N. Mitra, A. Surolia, *Biochim. Biophys. Acta* **2001**, *1527*, 102-111.

[53] D. M. Salunke, M. J. Swamy, M. Islam Khan, S. C. Mande, A. Surolia, M. Vijayan, *J. Biol. Chem.* **1985**, *260*, 13576-13579.

[54] M. Decastel, H. De Boeck, Y. Goussault, C. K. De Bruyne, F. G. Loontiens, J.-P. Frénoy, *Arch. Biochem. Biophys.* **1985**, *240*, 811-819.

[55] D. M. Salunke, M. Islam Khan, A. Surolia, M. Vijayan, *J. Mol. Biol.* **1982**, *154*, 177-178.

[56] D. M. Salunke, M. Islam Khan, A. Surolia, M. Vijayan, *FEBS Lett.* **1983**, *156*, 127-129.

[57] R. Ravishankar, K. Suguna, A. Surolia, M. Vijayan, *Acta Crystallogr.* **1999**, *D55*, 1375-1382.

[58] R. Ravishankar, M. Ravindran, K. Suguna, A. Surolia, M. Vijayan, *Curr. Sci.* **1997**, *72*, 855-861.

[59] R. Ravishankar, M. Ravindran, K. Suguna, A. Surolia, M. Vijayan, *Curr. Sci.* **1999**, *76*, 1393-1404.

[60] J. V. Pratap, G. M. Bradbrook, G. B. Reddy, A. Surolia, J. Raftery, J. R. Helliwell, M. Vijayan, *Acta Crystallogr.* **2001**, *D57*, 1584-1594.

[61] N. Sharon, *TIBS* **1993**, *18*, 221-226.

[62] S. Elgavish, B. Shaanan, *J. Mol. Biol.* **1998**, *277*, 917-932.

[63] V. Sharma, M. Vijayan, A. Surolia, *J. Biol. Chem.* **1996**, *271*, 21209-21213.

[64] P. Adhikari, K. Bachhawat-Sikder, C. J. Thomas, R. Ravishankar, A. A. Jeyaprakash, V. Sharma, M. Vijayan, A. Surolia, *J. Biol. Chem.* **2001**, *276*, 40734-40739.

[65] K. J. Neurohr, N. M. Young, I. C. P. Smith, H. H. Mantsch, *Biochemistry* **1981**, *20*, 3499-3504.

[66] K. J. Neurohr, H. H. Mantsch, N. M. Young, D. R. Bundle, *Biochemistry* **1982**, *21*, 498-503.

[67] M. Decastel, A.-T. Tran, J.-P. Frénoy, *Biochem. Biophys. Res. Comm.* **1982**, *106*, 638-643.

[68] H. De Boeck, K. L. Matta, M. Claeyssens, N. Sharon, F. G. Loontiens, *Eur. J. Biochem.* **1983**, *131*, 453-460.

[69] M. M. Prabu, R. Sankaranarayanan, K. D. Puri, V. Sharma, A. Surolia, M. Vijayan, K. Suguna, *J. Mol. Biol.* **1998**, *276*, 787-796.

[70] A. Cooper, *Curr. Opin. Chem. Biol.* **1999**, *3*, 557-563.

[71] Y.-Z. Liang, Z.-C. Li, F.-M. Li, *J. Colloid Interface Sci.* **2000**, *224*, 84-90.

[72] E. E. Dormidontova, *Macromolecules* **2002**, *35*, 987-1001.

[73] J. J. Christensen, H. L.D., R. M. Izatt, in *Handbook of Proton Ionization Heats and Related Thermodynamic Quantities*, John Wiley, New York, **1976**.

[74] U. Rungsardthong, *Doctoral Thesis*, University of Nottingham, U.K., **2001**.

[75] L. R. Olsen, A. Dessen, D. Gupta, S. Sabesan, J. C. Sacchettini, C. F. Brewer, *Biochemistry* **1997**, *36*, 15073-15080.

[76] J. C. Sacchettini, L. G. Baum, C. F. Brewer, *Biochemistry* **2001**, *40*, 3009-3015.

[77] M. Kanai, K. H. Mortell, L. L. Kiessling, *J. Am. Chem. Soc.* **1997**, *119*, 9931-9932.

[78] T. Wiseman, S. Williston, J. F. Brandts, L. N. Lin, *Anal. Biochem.* **1989**, *179*, 131-137.

[79] T. Govender, T. Ehtezazi, S. Stolnik, L. Illum, S. S. Davis, *Pharm. Res.* **1999**, *16*, 1125-1131.

APPENDIX I

TABLE 1. CRYSTAL DATA AND STRUCTURE REFINEMENT FOR ACGLCEMA (1A).

Identification code	AcGlcEMA	
Empirical formula	C20 H28 O12	
Formula weight	460.42	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Orthorhombic	
Space group	$P2_12_12_1$ (No. 19)	
Unit cell dimensions	a = 8.442(2) Å	α= 90°
	b = 12.466(4) Å	β= 90°
	c = 21.904(7) Å	γ = 90°
Volume	2305(1) Å^3	
Z	4	
Density (calculated)	1.327 g/cm^3	
Absorption coefficient	0.111 mm^{-1}	
F(000)	976	
Crystal size	0.1 × 0.7 × 1.2 mm^3	
θ range for data collection	1.86 to 29.0°.	
Index ranges	$-11 \leq h \leq 11, -16 \leq k \leq 16, -27 \leq l \leq 29$	
Reflections collected	28228	
Independent reflections	6082 [R(int) = 0.0479]	
Reflections with I>2σ(I)	5679	
Completeness to θ = 29.00°	99.7 %	
Absorption correction	None	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	6082 / 0 / 307	
Largest final shift/e.s.d. ratio	0.021	
Goodness-of-fit on F^2	1.051	
Final R indices [I>2σ(I)]	R1 = 0.0365, wR2 = 0.0938	
R indices (all data)	R1 = 0.0404, wR2 = 0.0971	
Absolute structure parameter	-0.1(6)	
Largest diff. peak and hole	0.326 and -0.192 e.Å^{-3}	

TABLE 2. ATOMIC COORDINATES ($\times 10^5$) AND EQUIVALENT ISOTROPIC DISPLACEMENT PARAMETERS ($Å^2 \times 10^4$) FOR ACGLCEMA. U(EQ) IS DEFINED AS ONE THIRD OF THE TRACE OF THE ORTHOGONALISED U_{IJ} TENSOR.

	x	y	z	U(eq)
O(1)	66194(11)	61041(8)	24748(5)	240(2)
O(2)	39079(11)	54306(8)	18033(4)	229(2)
O(3)	13277(11)	46898(7)	25450(4)	212(2)
O(4)	8428(12)	60440(8)	35890(5)	254(2)
O(5)	51166(11)	62381(7)	33364(4)	219(2)
O(6)	42843(13)	56365(8)	45196(5)	275(2)
O(8)	58951(13)	78650(8)	16365(5)	266(2)
O(9)	59057(18)	71625(15)	6945(6)	526(4)
O(13)	51326(16)	38315(9)	17123(5)	382(3)
O(15)	1913(12)	57142(8)	18126(5)	279(2)
O(17)	6004(16)	46759(13)	42588(7)	486(3)
O(19)	49410(30)	64746(13)	53805(7)	755(7)
C(1)	50882(15)	62774(10)	26854(6)	205(2)
C(2)	40118(14)	53788(10)	24604(6)	195(2)
C(3)	23506(14)	55721(10)	27095(6)	187(2)
C(4)	23827(15)	56685(10)	34018(6)	198(2)
C(5)	36050(15)	65185(10)	35870(6)	209(2)
C(6)	38146(18)	66582(11)	42664(7)	256(3)
C(7)	75315(15)	70757(11)	24041(6)	228(3)
C(8)	74808(16)	74506(12)	17498(6)	255(3)
C(9)	52095(19)	76140(13)	10982(6)	313(3)
C(10)	35060(20)	79499(13)	10784(7)	329(3)
C(11)	28200(20)	85410(15)	16094(9)	401(4)
C(12)	26720(30)	76867(18)	5843(9)	450(4)
C(13)	44977(16)	45944(11)	14805(7)	255(3)
C(14)	42500(30)	47516(16)	8088(8)	427(4)
C(15)	2865(14)	48696(11)	20778(6)	219(2)
C(16)	-6898(18)	38952(13)	19405(7)	305(3)
C(17)	773(17)	54798(13)	40337(6)	298(3)
C(18)	-14580(20)	60146(17)	41915(8)	404(4)
C(19)	48230(20)	56550(14)	50942(7)	366(3)
C(20)	52670(30)	45544(15)	53117(8)	436(4)

TABLE 3. BOND LENGTHS [Å] AND ANGLES [°] FOR ACGLCEMA.

O(1)-C(1)	1.3895(16)	O(15)-C(15)	1.2051(18)
O(1)-C(7)	1.4435(16)	O(17)-C(17)	1.201(2)
O(2)-C(13)	1.3545(17)	O(19)-C(19)	1.203(2)
O(2)-C(2)	1.4434(16)	C(1)-C(2)	1.5243(18)
O(3)-C(15)	1.3675(16)	C(2)-C(3)	1.5239(17)
O(3)-C(3)	1.4441(15)	C(3)-C(4)	1.5213(18)
O(4)-C(17)	1.3642(18)	C(4)-C(5)	1.5337(18)
O(4)-C(4)	1.4412(16)	C(5)-C(6)	1.509(2)
O(5)-C(1)	1.4269(16)	C(7)-C(8)	1.508(2)
O(5)-C(5)	1.4325(17)	C(9)-C(10)	1.498(2)
O(6)-C(19)	1.3385(18)	C(10)-C(12)	1.332(3)
O(6)-C(6)	1.4446(18)	C(10)-C(11)	1.494(3)
O(8)-C(9)	1.3504(18)	C(13)-C(14)	1.499(2)
O(8)-C(8)	1.4562(18)	C(15)-C(16)	1.498(2)
O(9)-C(9)	1.202(2)	C(17)-C(18)	1.498(2)
O(13)-C(13)	1.2040(19)	C(19)-C(20)	1.500(2)
C(1)-O(1)-C(7)	113.66(10)	O(6)-C(6)-C(5)	108.00(11)
C(13)-O(2)-C(2)	117.63(10)	O(1)-C(7)-C(8)	110.29(11)
C(15)-O(3)-C(3)	116.51(10)	O(8)-C(8)-C(7)	107.33(11)
C(17)-O(4)-C(4)	117.58(11)	O(9)-C(9)-O(8)	122.78(16)
C(1)-O(5)-C(5)	111.08(10)	O(9)-C(9)-C(10)	125.38(15)
C(19)-O(6)-C(6)	116.04(12)	O(8)-C(9)-C(10)	111.83(14)
C(9)-O(8)-C(8)	117.42(12)	C(12)-C(10)-C(11)	123.28(18)
O(1)-C(1)-O(5)	108.11(10)	C(12)-C(10)-C(9)	117.52(18)
O(1)-C(1)-C(2)	109.45(10)	C(11)-C(10)-C(9)	119.19(14)
O(5)-C(1)-C(2)	107.94(10)	O(13)-C(13)-O(2)	123.46(13)
O(2)-C(2)-C(3)	107.10(10)	O(13)-C(13)-C(14)	125.41(14)
O(2)-C(2)-C(1)	109.01(10)	O(2)-C(13)-C(14)	111.13(13)
C(3)-C(2)-C(1)	108.46(10)	O(15)-C(15)-O(3)	123.14(12)
O(3)-C(3)-C(4)	108.64(10)	O(15)-C(15)-C(16)	125.10(12)
O(3)-C(3)-C(2)	109.91(10)	O(3)-C(15)-C(16)	111.77(12)
C(4)-C(3)-C(2)	110.67(10)	O(17)-C(17)-O(4)	123.31(14)
O(4)-C(4)-C(3)	107.05(10)	O(17)-C(17)-C(18)	126.49(15)
O(4)-C(4)-C(5)	107.89(11)	O(4)-C(17)-C(18)	110.19(14)
C(3)-C(4)-C(5)	109.29(10)	O(19)-C(19)-O(6)	122.22(16)
O(5)-C(5)-C(6)	107.56(11)	O(19)-C(19)-C(20)	126.19(15)
O(5)-C(5)-C(4)	109.23(10)	O(6)-C(19)-C(20)	111.58(14)
C(6)-C(5)-C(4)	114.81(11)		

TABLE 4. ANISOTROPIC DISPLACEMENT PARAMETERS ($Å^2 \times 10^3$) FOR ACGLCEMA. THE ANISOTROPIC DISPLACEMENT FACTOR EXPONENT TAKES THE FORM: $-2\pi^4[h^2 a^{*2}U_{11} + ... + 2 hka^* b^* U_{12}]$

	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O(1)	207(4)	177(4)	335(5)	16(4)	28(4)	4(3)
O(2)	272(4)	216(4)	198(4)	11(4)	6(3)	54(4)
O(3)	228(4)	185(4)	224(4)	9(3)	-40(3)	-32(3)
O(4)	233(4)	268(5)	261(5)	-9(4)	36(4)	0(4)
O(5)	214(4)	209(4)	235(4)	16(3)	-19(4)	-1(4)
O(6)	360(5)	239(5)	228(5)	18(4)	-57(4)	-18(4)
O(8)	309(5)	243(5)	246(5)	-1(4)	-36(4)	11(4)
O(9)	481(7)	846(11)	251(6)	-119(6)	64(5)	-101(8)
O(13)	565(7)	230(5)	350(6)	-6(4)	74(6)	122(5)
O(15)	313(5)	275(5)	248(5)	10(4)	-52(4)	48(4)
O(17)	387(6)	606(9)	466(7)	257(7)	58(5)	-39(6)
O(19)	1473(19)	476(8)	314(7)	-134(6)	-309(9)	289(11)
C(1)	202(5)	182(6)	230(6)	18(4)	-4(5)	10(5)
C(2)	210(5)	182(5)	194(5)	9(4)	-4(4)	12(4)
C(3)	205(5)	150(5)	206(5)	9(4)	-15(4)	-17(4)
C(4)	211(5)	180(5)	204(5)	-4(5)	0(4)	0(5)
C(5)	230(5)	175(5)	221(6)	4(5)	-13(5)	-6(4)
C(6)	319(7)	214(6)	235(6)	-25(5)	-24(5)	-10(5)
C(7)	205(5)	210(6)	268(6)	-10(5)	3(5)	-22(5)
C(8)	259(6)	245(6)	263(6)	-7(5)	37(5)	-9(5)
C(9)	369(7)	355(8)	216(6)	68(5)	-9(5)	-114(6)
C(10)	369(7)	312(7)	306(8)	120(6)	-64(6)	-89(6)
C(11)	371(8)	412(9)	422(9)	37(8)	-98(7)	84(7)
C(12)	485(10)	464(10)	402(9)	60(8)	-136(8)	-83(8)
C(13)	270(6)	227(6)	268(6)	-22(5)	49(5)	-7(5)
C(14)	622(11)	417(9)	241(7)	-35(7)	45(8)	86(9)
C(15)	188(5)	267(6)	201(6)	-40(5)	0(4)	19(5)
C(16)	284(6)	326(7)	305(7)	-37(6)	-43(6)	-59(6)
C(17)	255(6)	416(8)	224(6)	-28(6)	5(5)	-96(6)
C(18)	270(7)	581(11)	360(9)	-99(8)	74(6)	-68(7)
C(19)	517(9)	383(8)	198(6)	-19(6)	-8(6)	89(8)
C(20)	662(12)	407(9)	238(7)	30(7)	-5(7)	147(9)

TABLE 5. HYDROGEN COORDINATES ($\times 10^4$) AND ISOTROPIC DISPLACEMENT PARAMETERS ($Å^2 \times 10^3$) FOR ACGLCEMA.

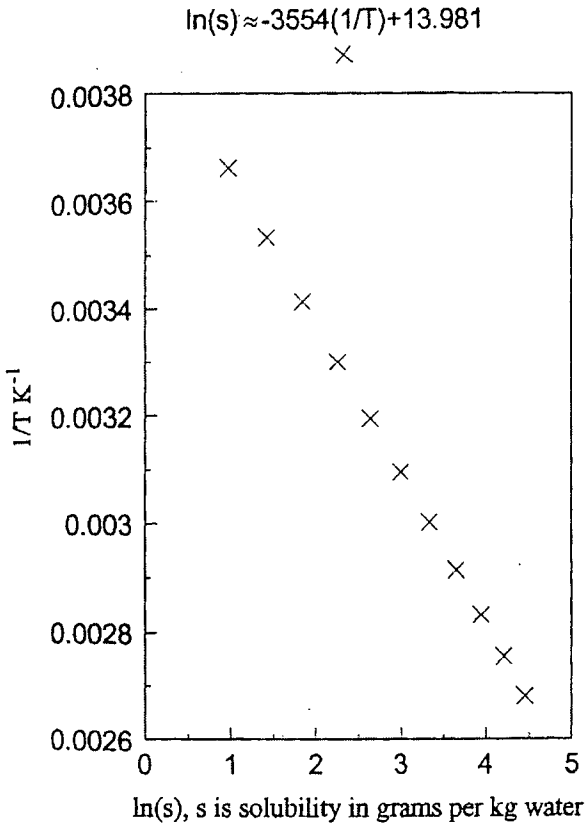
	x	y	z	U(iso)
H(1)	4687	6989	2543	27
H(2)	4422	4664	2594	25
H(3)	1919	6250	2531	24
H(4)	2632	4962	3595	26
H(5)	3273	7223	3410	27
H(6A)	4638	7205	4349	33
H(6B)	2810	6900	4454	33
H(7A)	8643	6941	2526	30
H(7B)	7097	7641	2674	30
H(8A)	8278	8021	1682	33
H(8B)	7714	6847	1470	33
H(111)	1688	8668	1541	54(4)
H(112)	3367	9230	1654	54(4)
H(113)	2961	8117	1982	54(4)
H(121)	3190(20)	7298(17)	297(10)	34(5)
H(122)	1560(30)	7890(20)	541(11)	51(6)
H(141)	5103	5196	641	90(6)
H(142)	3229	5106	739	90(6)
H(143)	4250	4051	605	90(6)
H(161)	-1817	4082	1955	58(4)
H(162)	-465	3340	2245	58(4)
H(163)	-426	3625	1533	58(4)
H(181)	-2017	5588	4500	66(4)
H(182)	-2116	6074	3824	66(4)
H(183)	-1245	6733	4354	66(4)
H(201)	5034	4491	5749	78(5)
H(202)	6400	4436	5243	78(5)
H(203)	4656	4017	5085	78(5)

TABLE 6. TORSION ANGLES [°] FOR ACGLCEMA.

C(7)-O(1)-C(1)-O(5)	-90.14(12)	O(4)-C(4)-C(5)-O(5)	-172.38(10)
C(7)-O(1)-C(1)-C(2)	152.53(11)	C(3)-C(4)-C(5)-O(5)	-56.32(13)
C(5)-O(5)-C(1)-O(1)	174.12(10)	O(4)-C(4)-C(5)-C(6)	66.73(14)
C(5)-O(5)-C(1)-C(2)	-67.58(12)	C(3)-C(4)-C(5)-C(6)	-177.21(11)
C(13)-O(2)-C(2)-C(3)	-127.07(11)	C(19)-O(6)-C(6)-C(5)	167.34(13)
C(13)-O(2)-C(2)-C(1)	115.77(12)	O(5)-C(5)-C(6)-O(6)	-63.42(13)
O(1)-C(1)-C(2)-O(2)	-65.04(12)	C(4)-C(5)-C(6)-O(6)	58.39(15)
O(5)-C(1)-C(2)-O(2)	177.52(9)	C(1)-O(1)-C(7)-C(8)	-96.81(13)
O(1)-C(1)-C(2)-C(3)	178.67(10)	C(9)-O(8)-C(8)-C(7)	-138.34(12)
O(5)-C(1)-C(2)-C(3)	61.23(12)	O(1)-C(7)-C(8)-O(8)	71.93(13)
C(15)-O(3)-C(3)-C(4)	135.20(11)	C(8)-O(8)-C(9)-O(9)	-7.5(2)
C(15)-O(3)-C(3)-C(2)	-103.57(12)	C(8)-O(8)-C(9)-C(10)	171.57(11)
O(2)-C(2)-C(3)-O(3)	66.99(12)	O(9)-C(9)-C(10)-C(12)	3.1(3)
C(1)-C(2)-C(3)-O(3)	-175.49(10)	O(8)-C(9)-C(10)-C(12)	-175.97(15)
O(2)-C(2)-C(3)-C(4)	-173.01(10)	O(9)-C(9)-C(10)-C(11)	-177.85(18)
C(1)-C(2)-C(3)-C(4)	-55.50(13)	O(8)-C(9)-C(10)-C(11)	3.1(2)
C(17)-O(4)-C(4)-C(3)	128.43(12)	C(2)-O(2)-C(13)-O(13)	-1.9(2)
C(17)-O(4)-C(4)-C(5)	-114.05(13)	C(2)-O(2)-C(13)-C(14)	178.02(13)
O(3)-C(3)-C(4)-O(4)	-69.67(12)	C(3)-O(3)-C(15)-O(15)	-0.64(18)
C(2)-C(3)-C(4)-O(4)	169.57(10)	C(3)-O(3)-C(15)-C(16)	178.99(11)
O(3)-C(3)-C(4)-C(5)	173.73(10)	C(4)-O(4)-C(17)-O(17)	-2.4(2)
C(2)-C(3)-C(4)-C(5)	52.98(13)	C(4)-O(4)-C(17)-C(18)	177.11(12)
C(1)-O(5)-C(5)-C(6)	-169.65(10)	C(6)-O(6)-C(19)-O(19)	-1.5(3)
C(1)-O(5)-C(5)-C(4)	65.13(13)	C(6)-O(6)-C(19)-C(20)	179.68(14)

APPENDIX II

SOLUBILITY DATA FOR DL-ASPARTIC ACID



Solubility curve for dl-aspartic acid[1].

From the curve reported above, the percentage supersaturation of the system, $\alpha\%$, can be determined as follows:

$$\alpha\% = \frac{s - s_{sat}}{s_{sat}} \times 100$$

where: s = amount dissolved in solution

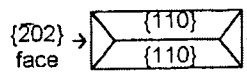
s_{sat} = amount dissolved in saturated solution

[1] S. J. Cooper, *Doctoral Thesis*, University of Bristol, U.K., **1993**, p. 144.

APPENDIX III

CRYSTAL HABIT OF DL-ASPARTIC ACID[1]

CRYSTALLOGRAPHIC DATA FOR HOMOGENEOUSLY GROWN DL-ASPARTIC ACID[2]

Space group	C2/c (number 15)
Unit cell parameters	a = 18.947A°, b = 7.433A° c = 9.184A°, β = 123.75°
Number of molecules per unit cell	8
Reflection conditions	h00: h = 2n, 0k0: k = 2n, 00l: l = 2n, h0l: h, l = 2n, 0kl: k = 2n, hk0: h + k = 2n, hkl: h + k = 2n
Typical crystal habit in aqueous solution	 other faces sometimes present: {200}, {020}, {002} and {111}
Nomenclature	Monoclinic crystal {h00}, {0k0}, {00l} = pinacoid faces {hk0} = prisms, {h0l} and {0kl} = domes {hkl} = pyramids forms parallel to *a* axis = clino forms parallel to *b* axis = ortho

The Cerius morphology module from MSI (Molecular Simulations Incorporated) predicts the external morphology of the crystalline material from the internal crystal structure. The Bravais Friedel Donnay Harker (BFDH) method and the Attachment Energy (AE) method for growth morphologies are used to determine morphologies in vacuum, since no account is taken of the interaction of the crystal faces with the solvent (water). The BFDH method uses the crystal lattice dimensions

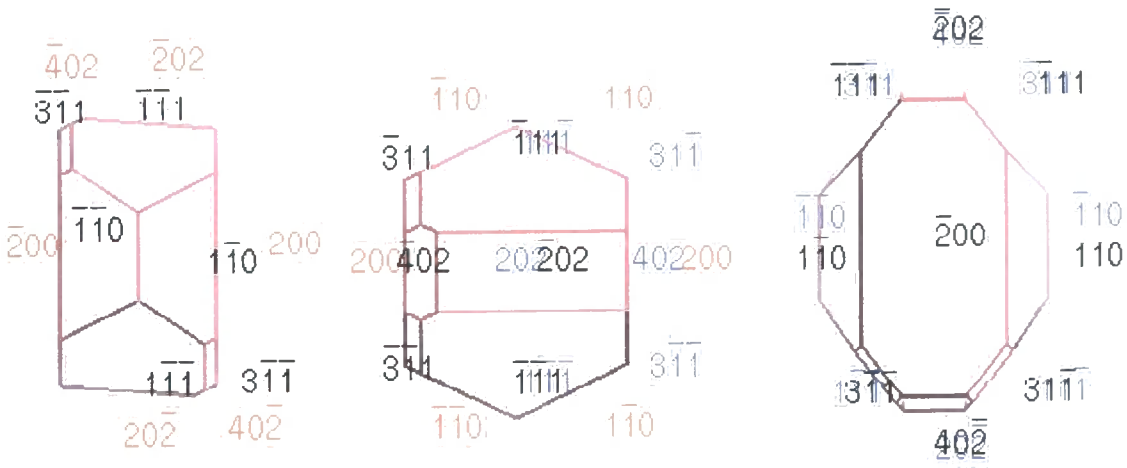
[1] S. Al-Mannai, *Master Dissertation*, University of Durham, U.K., **2001**, 75-79.
[2] S. J. Cooper, *Doctoral Thesis*, University of Bristol, U.K., **1993**, p. 144.

and symmetry to generate a list of possible growth faces and their relative growth rates. The AE method relies on calculations of the energy released when a growth slice is added to a growing plane. The force field COMPASS was used to calculate the crystal and attachment energies. Due to the consideration of the energetics of the system, the AE method gives more accurate predictions.

Comparing the typical aqueous morphology with the predicted vacuum morphologies obtained by the BFDH and AE methods, it is clear that both the methods overestimate the importance of the faces $\{\bar{2}02\}$, $\{200\}$ and $\{\bar{3}11\}$ faces. Of the two, the AE method is slightly more accurate, because the $\{\bar{2}02\}$ and $\{\bar{3}11\}$ faces have smaller areas.

AREAS OF DL-ASPARTIC ACID CRYSTAL FACES CALCULATED BY BFDH METHOD

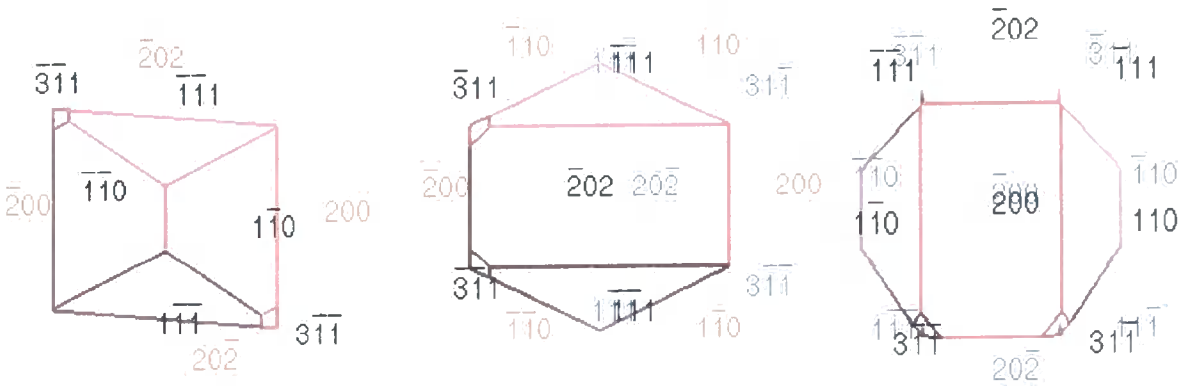
CRYSTAL FACE	CRYSTAL FACE AREA
$\{\bar{1}\ 1\ 1\}$	7.18
$\{1\ 1\ 0\}$	6.48
$\{\bar{2}\ 0\ 2\}$	4.22
$\{2\ 0\ 0\}$	16.68
$\{\bar{3}\ 1\ 1\}$	0.44



Three perpendicular views of the crystal morphology predicted for dl-aspartic acid by BFDH method. Black: faces in front; blue: faces under; red: faces perpendicular to the view.

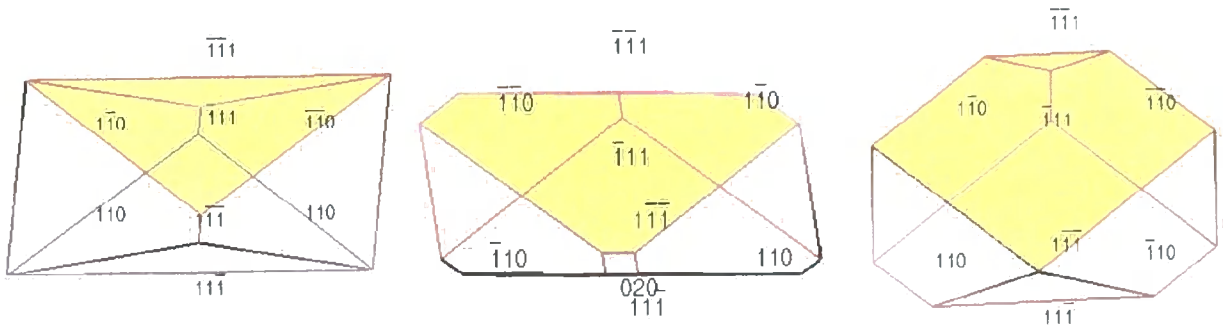
AREAS OF DL-ASPARTIC ACID CRYSTAL FACES CALCULATED BY AE METHOD

CRYSTAL FACE	CRYSTAL FACE AREA
$\{\bar{1}\ 1\ 1\}$	4.77
$\{1\ 1\ 0\}$	7.74
$\{\bar{2}\ 0\ 2\}$	12.95
$\{2\ 0\ 0\}$	11.60
$\{\bar{3}\ 1\ 1\}$	0.21

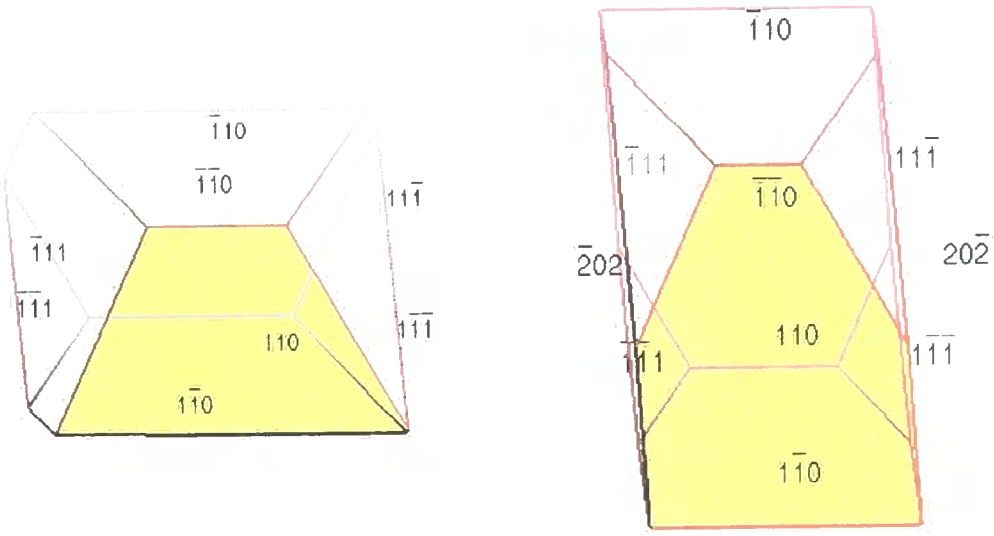


Three perpendicular views of the crystal morphology predicted for dl-aspartic acid by AE method. Black: faces in front; blue: faces under; red: faces perpendicular to the view.

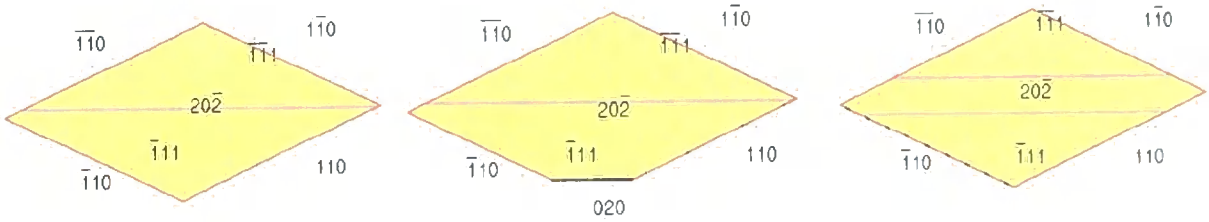
The following figures for crystal morphologies are produced from the Cerius[2] software by varying the sizes of the crystal faces present. In particular, the face growing beneath the monolayer film is made larger, according to the experimental observation.



Three different shaped $\{\bar{1}\ 1\ 1\}$ crystal faces (yellow).



Two different shaped {110} crystals faces (yellow).



Three different shaped { $\bar{2}$02} crystals faces (yellow).